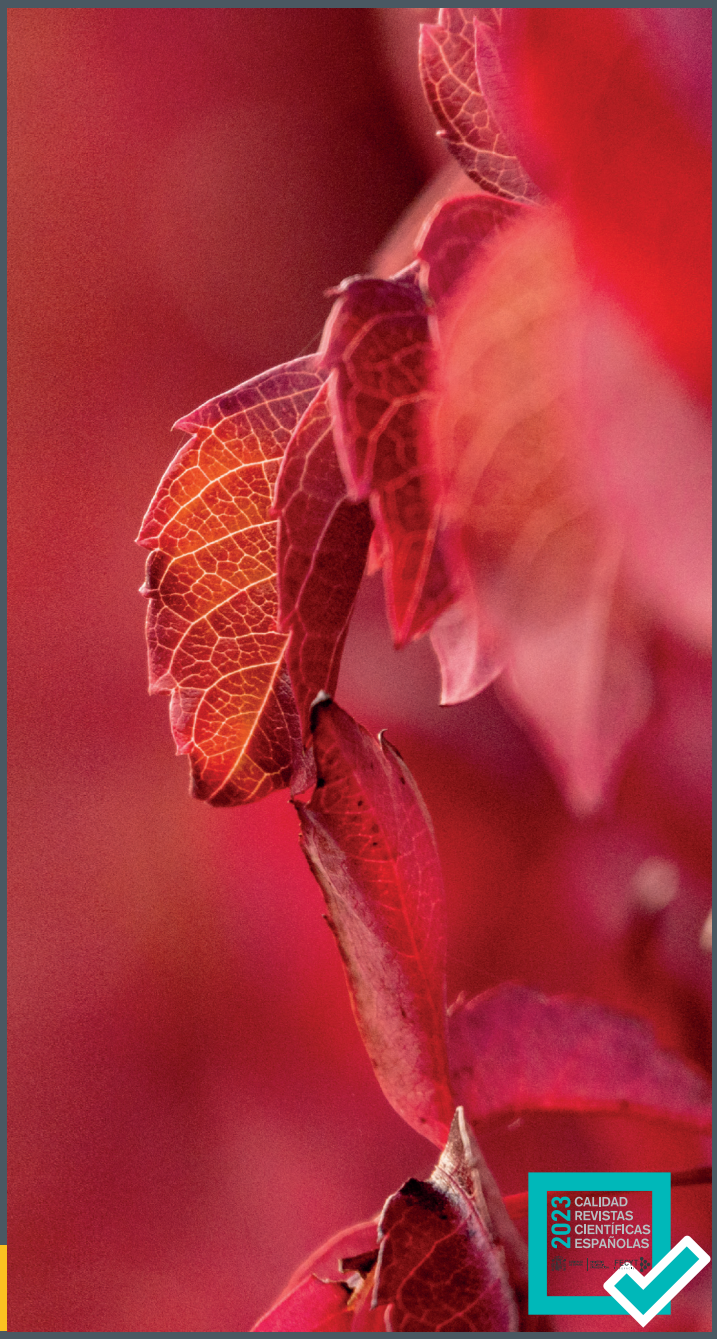
revista de EDUCACIÓN

Nº 406 OCTUBRE-DICIEMBRE 2024



2023 CALIDAD REVISTAS CIENTÍFICAS ESPAÑOLAS

GOBIERNO DE ESPAÑA
MINISTERIO DE EDUCACIÓN, FORMACIÓN PROFESIONAL Y DEPORTES

revista de
EDUCACIÓN

Nº 406 OCTUBRE-DICIEMBRE 2024

revista de EDUCACIÓN

Nº 406 Octubre-Diciembre 2024
Revista trimestral
Fecha de inicio: 1952

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

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: https://www.libreria.educacion.gob.es/
Catálogo general de publicaciones oficiales de la Administración General del Estado: https://cpage.mpr.gob.es/

Edición: 2024
NIPO pdf: 164-24-063-9
NIPO ibd: 164-24-061-8
NIPO html: 164-24-062-3
ISSN línea: 1988-592X
ISSN papel: 0034-8082
Depósito Legal: M.57/1958
Imagen de cubierta: Annie Spratt en Unplash
Diseño de la portada: Dinarte S.L.
Maqueta: Glaux Publicaciones Académicas S.L.U.

CONSEJO DE DIRECCIÓN

PRESIDENCIA

Abelardo de la Rosa
Secretario de Estado de Educación

Clara Sanz López
Secretaria General de Formación Profesional

VOCALES

Santiago Antonio Roura Gómez
Subsecretario de Educación, Formación Profesional y deportes

Mónica Domínguez García
Directora General de Evaluación y Cooperación Territorial

Carmen Tovar Sánchez
Directora del Instituto Nacional de Evaluación Educativa

Clementa López Fernández
Subdirectora General de Atención al Ciudadano, Documentación y Publicaciones

Francisco Javier Amaya Flores
Subdirectora de Cooperación Territorial e Innovación Educativa

Helena Ramos García
Subdirector General de Ordenación Académica

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: Sabrina Gallego Verdi
Colaboradores: Ruth Martín Escanilla y Alberto Díaz-Roncero Canales

ASESORES CIENTÍFICOS

Internacional

Aaron Benavot (State University of New York, SUNY-Albany); Abdeljalilb 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); 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 Muñiz (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); Jesús García Laborda (Universidad de Alcalá).

Índice

Editorial

José Luis Gaviria Soto:
Presentación: La validación de instrumentos de medida a examen 1

Réplicas

José A. Martínez García: Crítica del análisis de la validez de constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GRS 2); réplica a Tourón et al. (2023) 7

Marta Tourón, Enrique Navarro-Asencio y Javier Tourón: Réplica a la "Crítica del análisis de la validez de constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GRS 2); réplica a Tourón et al. (2023)"..... 35

Fernando Martínez Abad y José Carlos Sánchez Prieto: Contribución al conocimiento en Investigación Educativa a partir de estudios de validación de escalas: una reflexión crítica sobre el trabajo de M. Tourón et al. (2023) y el análisis metodológico de Martínez-García (2024) 65

Investigaciones

Muammer Maral: Un análisis bibliométrico sobre la eficiencia en la educación..... 85

Jing Xu: La efectividad de la realidad virtual basada en video en la educación artística para el desarrollo de la inteligencia social y emocional en estudiantes de secundaria 129

Santiago Vicente Martín, Marta Ramos Baz, Mª. Mercedes Rodríguez Sánchez, Beatriz Sánchez-Barbero y Rosario Sánchez Fernández: Problemas aritméticos verbales en los libros castellano-leoneses y andaluces 157

María Teresa Pozo Llorente, María de Fátima Poza Vilches y José Manuel Aguayo Moral: La Formación Profesional Dual en los centros educativos andaluces desde la perspectiva del profesorado..... 185

Celia Camili Trujillo, Lorena Gil Pastor, Sara Serrano Díaz y Luis Aimá González: Formación para el desarrollo profesional docente en escuelas: revisión sistemática 219

Inmaculada Sureda-García, Rafael Jiménez, Albert Sesé y Francesca Salvà-Mut: Validación española del Instrumento de Vinculación Estudiantil en Formación Profesional..... 249

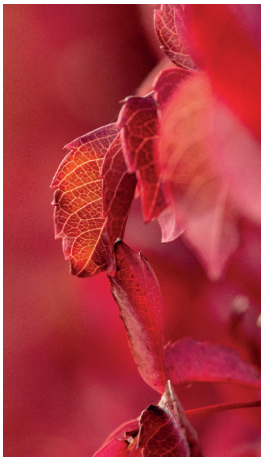
María Reneses, María Riberas-Gutiérrez y Nereida Bueno-Guerra: ¿Quién comparte fake news? Consumo y distribución de información entre adolescentes y su relación con el discurso de odio 281

Olatz Arce-Larrory, Erlantz Velasco y Iker Sáez: Hábitos de vida saludables en educación primaria: revisión sistemática..... 309

Ana Ayuste González y Jaume Trilla Bernet: Pensamiento Crítico y Pedagogía Crítica: coincidencias y complementariedades 345

Silvia Abad Merino, Fernando Macías-Aranda, Olga Serradell Pumareda y Ainhoa Flecha: La experiencia académica del alumnado gitano en la universidad..... 369

Reseñas..... 407



Editorial

Presentación: La validación de instrumentos de medida a examen

Presentation: Measurement instruments validation process under scrutiny

José Luis Gaviria Soto
Editor Jefe
Revista de Educación

En el presente número, la Revista de Educación lleva a cabo una iniciativa interesante. Hemos compilado un grupo de trabajos alrededor del análisis crítico de un artículo recientemente publicado en esta revista.

El propósito de cualquier revista académica consiste en servir de medio de comunicación para la comunidad científica a la que sirve. Los medios tecnológicos han posibilitado que en la actualidad la comunicación sea bidireccional y prácticamente inmediata. Sin embargo, las revistas científicas todavía cumplen el papel de favorecer la comunicación pausada y reflexiva de las novedades que en el ámbito científico de referencia se producen. Pero en ocasiones los académicos necesitan ir un paso más allá de la recepción de información, y tienen que poder pasar a un intercambio de ideas, a una contraposición de argumentos que permita clarificar los términos usados y depurar las metodologías aceptadas. En este caso el tema sobre el que se produce este intercambio enriquecedor es el de la validación de los instrumentos de medida utilizados en el ámbito educativo.

La validación es un proceso fundamental para garantizar que la información de las variables utilizadas tiene un contenido claro y que refleja acertadamente la realidad que los investigadores quieren medir.

En la ciencia tradicional la cuestión de la relación entre el instrumento de medida y la realidad medida es una cuestión técnica que, por lo general, tiene como mayor preocupación la precisión, es decir, la correspondencia entre el número producido y la realidad medida. Pero incluso en los ámbitos científicos más tradicionales, se ha llegado a un

punto en el que la pura comparación directa de la realidad empírica con el instrumento de medida es imposible, y sólo mediante procedimientos indirectos es posible llevar a buen fin la medida. Esto plantea el problema de si lo que se está midiendo responde realmente al objetivo original.

En las ciencias sociales en general, y en psicología y educación en particular, ese problema ha estado presente desde sus inicios como disciplinas científicas. ¿Estamos midiendo realmente lo que decimos? ¿Hay una verdadera relación entre la magnitud que definimos teóricamente y la escala que producimos durante la medida? ¿Es el constructo que hemos definido teóricamente el que está determinando los valores de la escala, o es el acto de medir el que está creando el constructo?

Todas estas son preocupaciones que vienen de lejos, y son las que han determinado los procedimientos que se han ido depurando hasta constituir un conjunto de técnicas que son utilizadas de manera estándar por los científicos sociales.

Pero ésta nunca ha sido una cuestión cerrada. La preocupación por la validez de los instrumentos utilizados en nuestro ámbito científico está muy justificada. No en vano tratamos prácticamente siempre de medir características de los individuos que no son directamente observables, y de las que sólo podemos obtener evidencia indirecta. Por eso hay tal proliferación de estudios de validación en la literatura científica de nuestras áreas. Y como indican Martínez Abad y Sánchez Prieto (2024) en este mismo número, la cantidad de artículos dedicados a la validación de constructos se ha disparado de manera casi exponencial en los últimos años.

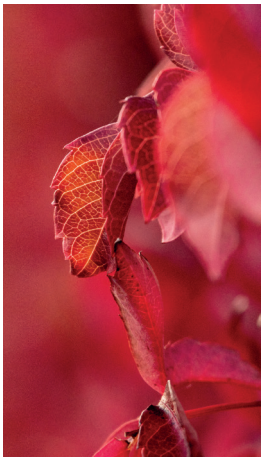
En el número 402 de nuestra revista se publicó el artículo 'Validez de Constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GSR 2), En España', (Tourón, Navarro-Asencio y Tourón, 2023). La particularidad de esta escala es que no trata de medir directamente la alta capacidad de los individuos, sino que se centra en la información proporcionada por los padres como fuente para hacer un screening previo. Es decir, que, si habitualmente diagnosticar a un sujeto respecto de cualquier rasgo latente supone llevar a cabo una medida a base de evidencias indirectas, aquí la situación es todavía más indirecta, pues tratamos de utilizar evidencias respecto de las percepciones de los padres de esos alumnos para tratar de identificar a los que pueden poseer altas capacidades.

Como es lógico, esto lleva a preguntarse cómo son las relaciones de causalidad entre los elementos del instrumento de medida y la realidad que queremos estudiar. ¿Es esa realidad, la alta capacidad, la que determina las percepciones de los padres, y estas a su vez se ven reflejadas en sus respuestas al GSR2, o son las preguntas del GSR 2 las que determinan la realidad que queremos definir?

Tomando este trabajo como estímulo para su reflexión, Martínez-García (2024) hace una reflexión crítica respecto a la naturaleza reflectiva o formativa del constructo propuesto. ¿Cuál es la verdadera dirección de la causalidad en la relación entre el instrumento y el constructo? Martínez-García se apoya en el análisis crítico del trabajo de Tourón et al. (2023) para considerar que posiblemente el estándar establecido en los trabajos de validación de constructo adolece de la falta de consideración de la posibilidad de que la dirección causal sea la inversa, es decir, que sea nuestro cuestionario el que está delimitando la realidad que medimos. La consideración de que con nuestro cuestionario tal vez estemos delimitando una realidad en cierto modo artificial, sin que esto quiera decir que no sea real.

La dirección de la Revista de Educación consideró que el tema era suficientemente importante como para dedicarle la atención que merece. Por eso en este número se publica una sección monográfica dedicada a la validación de constructo. Es un monográfico en cierto modo especial, puesto que el artículo original que ha dado lugar a este intercambio académico estaba publicado en el número 402. En este número, el 406, se incluye el análisis crítico de Martínez-García (2024), una contribución sobre el mismo tema de Martínez-Abad y Sánchez Prieto (2024) y la réplica de los autores del trabajo original, Tourón, Navarro-Asencio y Tourón (2024). Este conjunto de cuatro artículos constituye un dossier centrado en algunas consideraciones metodológicas sobre la validación de constructo que tiene el atractivo de estar centrado alrededor de un ejemplo concreto y sobre un constructo muy actual.

Revista de Educación considera que se trata de un intercambio académico de alto nivel, que puede ayudar a mejorar nuestras prácticas metodológicas. Desde aquí queremos agradecer la generosidad de los participantes en este intercambio académico que sabemos que será muy provechoso para quienes quieran abordar en adelante la validación de instrumentos de medida en nuestro ámbito.



Réplicas

Crítica del análisis de la validez de constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GRS 2); réplica a Tourón et al. (2023)

Criticism of the analysis of Construct Validity of the Gifted Rating Scales (GRS 2) Parent Form in Spain; a reply to Tourón et al. (2023)

https://doi.org/10.4438/1988-592X-RE-2024-406-649

José A. Martínez García
https://orcid.org/0000-0003-2131-9101
Universidad Politécnica de Cartagena

Resumen

El objetivo de este trabajo es reexaminar los métodos de análisis de la validez de constructo de la escala de detección de alumnos con altas capacidades para padres (GRS 2) en España, realizados por Tourón et al. (2023). Para ello, partimos de la propuesta de estos autores, criticando de forma constructiva algunos de sus procedimientos, y planteando alternativas de modelización y análisis. De este modo, se emplea un enfoque eminentemente didáctico, basado fundamentalmente en la literatura sobre modelos de ecuaciones estructurales y psicometría. La distinción entre modelos reflectivos y formativos, la dirección de causalidad y las ecuaciones subyacentes, la idoneidad de analizar la validez de constructo en una red nomológica yendo más allá del análisis factorial, y el test del modelo estructural usando la chi-cuadrado, son los principales ejes sobre los que se articula la crítica. Asimismo, se proponen diferentes opciones de modelización que son congruentes con líneas de investigación en medición de altas capacidades, y que podrán servir de base para seguir avanzando en esta disciplina en el futuro.

Palabras clave: escala de detección, altas capacidades, validez de constructo, análisis factorial confirmatorio, ecuaciones estructurales, modelos formativos.

Abstract

The objective of this study is to reexamine the methods used to analyze the construct validity of the Gifted Rating Scales (GRS 2) Parent Form in Spain, originally conducted by Tourón et al. (2023). To achieve this, we build upon the proposals of these authors, offering constructive criticism of some of their procedures and suggesting alternative modeling and analysis methods. Our approach is primarily didactic, drawing extensively from the literature on structural equation modeling and psychometrics. Key elements of our critique include the distinction between reflective and formative models, the direction of causality and the underlying equations, the appropriateness of analyzing construct validity within a nomological network beyond factor analysis, and the use of chi-square testing for the structural model. Additionally, we propose various modeling options that align with current research trends in the measurement of high abilities, providing a foundation for future advancements in this discipline.

Keywords: gifted rating scales, high ability, construct validity, confirmatory factor analysis, structural equation modeling, formative models.

Introducción

La investigación de Tourón et al. (2023) analizó la validez de constructo de la Escala de Padres de las GRS 2 (*Gifted Rating Scales*) en España, sobre una muestra final de 1109 padres y madres de niños y adolescentes (entre 4 y 18 años) con altas capacidades. Esta escala, basada en un modelo original de 3 dimensiones (capacidades cognitivas, capacidades creativas y artísticas, y habilidades socioemocionales) consta de 20 ítems repartidos en ellas. Tourón et al. (2023) adaptaron el cuestionario al español y procedieron al estudio de validación, obteniendo al final 4 factores: capacidad cognitiva, capacidad creativa, habilidades sociales y control emocional, agrupados dos a dos en sendos factores de orden superior: capacidad cognitivo-creativa y habilidades socio-emocionales, respectivamente.

Sin embargo, varios de los procedimientos metodológicos y de análisis de resultados descritos por Tourón et al. (2023) son cuestionables. Tanto el modelo factorial propuesto, basado en una visión reflectiva de la medición de las altas capacidades, como los índices de ajuste empleados para validar la estructura dimensional, de carácter aproximado, limitan seriamente la interpretación de los resultados.

Recibido: 14/12/2023 Aceptado: 01/06/2024

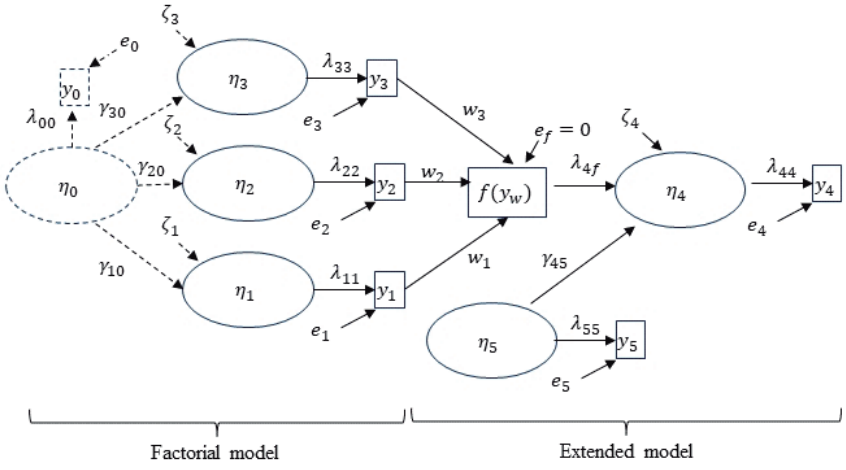
El objetivo de este trabajo es reexaminar los métodos de análisis de la escala GRS 2. Para ello, partimos de la propuesta de Tourón et al. (2023), criticando de forma constructiva algunos de sus procedimientos, y planteando alternativas de modelización y análisis. Así, nuestra propuesta no solo discute la idoneidad del enfoque de Tourón et al. (2023), sino que puede servir como estímulo para replantear algunos de los procedimientos de modelización y análisis que se realizan en la investigación sobre altas capacidades.

Para conseguir tal fin, vamos a emplear un enfoque eminentemente didáctico, basado fundamentalmente en la literatura sobre modelos de ecuaciones estructurales y psicometría. Dividiremos las críticas en varios puntos fundamentales a desarrollar en los siguientes apartados.

Modelización reflectiva frente a formativa

Vamos a considerar el modelo presentado en la Figura I como base. En ese modelo, hay un vector de variables latentes $\boldsymbol{\eta}$ cuyo valor verdadero (*true score*) es desconocido, pero que podemos intentar conocerlo a través de la relación de estas variables latentes con un vector de indicadores observables $\boldsymbol{y}$.

FIGURA I. Modelo factorial y modelo extendido



Recibido: 14/12/2023 Aceptado: 01/06/2024

La parte llamada "factorial" pretende representar la conceptualización inicial de Tourón et al. (2023); existe un modelo multidimensional de las altas capacidades: capacidades cognitivas, capacidades creativas y artísticas, y habilidades socioemocionales. En la Figura I este modelo está representado por las variables latentes η_1, η_2, η_3 y los indicadores observables (ítems) y_1, y_2, y_3. Recordemos que, según MacCallum y Austin (2000) las variables latentes son constructos hipotéticos que no pueden ser directamente medidos. Es decir, se reflejan a través de una serie de indicadores observables que, como revela Bollen (2002), son localmente independientes y definen el valor esperado de la variable latente, la cual es una función no determinista de esos indicadores.

En aras de simplificar, la Figura I solo contiene un indicador por variable latente, en lugar de los 6, 7 y 7, respectivamente, del modelo GRS 2. Esta divergencia no afecta a nuestra exposición inicial, aunque será importante más adelante. La relación entre las variables latentes y sus indicadores es causal, en línea con la teoría clásica de los test: la variación en la variable latente η_1 se manifiesta en una variación en el indicador observable, escalada con el parámetro λ_{11}, a lo que hay que añadir un error aleatorio e_1. Si los datos se parametrizan como desviaciones en torno a la media de los indicadores observables (algo habitual en programas informáticos especializados en este tipo de análisis, y que no afecta a las estructuras de covarianza), las ecuaciones que relacionan variables latentes e indicadores son tan sencillas como (1), representando la primera de las dimensiones:

$$y_1 = \lambda_{11}\eta_1 + e_1 \qquad (1)$$

Este tipo de estructura causal entre variables latentes e indicadores se denomina reflectiva (Bollen y Diamantopoulos, 2017), y los indicadores son manifestaciones (efectos) de la variable latente. Variaciones en el rasgo o factor latente producen variaciones en el indicador observable, en función de la elección (o estimación, según el caso) del factor de escala λ y del error e. La variable latente se escala con uno de sus indicadores (en este caso con el único indicador que tiene), fijando su valor (normalmente a la unidad, pero no necesariamente siempre así), lo que indica que variaciones unitarias en la variable latente se reflejan en variaciones también unitarias del indicador observable (más fluctuaciones aleatorias).

Recibido: 14/12/2023 Aceptado: 01/06/2024

La parte "factorial" de la Figura I la completa la variable latente η_0, que puede representar varios estados del mundo real. En primer lugar, se puede ver como una variable de orden superior a la estructura multidimensional, es decir, como una variable que representa una cualidad que no se puede medir pero que se manifiesta causalmente a través de dimensiones, las cuales sí se pueden medir con indicadores observables. Este tipo de conceptualización es empleada en los estudios sobre escalas GRS (ej. Sofologi et al., 2022). En segundo lugar, se puede interpretar como una variable latente que sí se puede medir a través de un indicador observable y_0, y que influye causalmente al resto de variables latentes. En la Figura I la relación entre variables latentes sigue una expresión similar a (1), que para la primera dimensión es (2):

$$\eta_1 = \gamma_{10}\eta_0 + \zeta_1 \quad (2)$$

Donde γ_{10} representa el efecto de η_0 sobre η_1, y ζ_1 es el error cometido en la determinación de η_1.

La conceptualización habitual en esta área de conocimiento es la primera, tal y como se muestra, por ejemplo, en Pfeiffer et al. (2008), donde se plantea una concepción multidimensional de las altas capacidades con un factor subyacente "g" o factor de capacidad general[1]. Hay que reconocer que Tourón et al. (2023) son muy cautos y no afirman claramente este hecho, pero subrepticiamente parecen hacerlo cuando calculan una "media total de la escala", como posteriormente explicaremos.

Por tanto, como hay un factor latente que no se puede medir directamente (no tenemos y_0), se plantea la medición de dimensiones que son manifestaciones de ese factor. Y esta distinción es crucial, porque la estimación del valor verdadero[2] de η_0 entonces debe hacerse a través de esta especificación (3):

[1] No es objetivo de este trabajo discutir acerca de las diferentes conceptualizaciones de las altas capacidades. Es evidente que la literatura es muy amplia al respecto y que hay discusión acerca de la definición, las políticas y las prácticas, como bien muestran Mcclain y Pfeiffer (2012). Simplemente vamos a considerar que se conciben las altas capacidades como una variable multidimensional o como una variable que no se puede medir directamente y que se manifiesta a través de diferentes dimensiones o rasgos.

[2] Es evidente que si los datos se parametrizan como desviaciones con respecto a la media el valor "verdadero" medio de cada variable latente es cero.

$$\eta_0 = \frac{\eta_1 - \zeta_1}{\gamma_{10}} = \frac{\left(\frac{y_1 - e_1}{\lambda_{11}}\right) - \zeta_1}{\gamma_{10}} \tag{3}$$

Así, se debe cumplir la igualdad (4):

$$\eta_0 = \frac{\left(\frac{y_1 - e_1}{\lambda_{11}}\right) - \zeta_1}{\gamma_{10}} = \frac{\left(\frac{y_2 - e_2}{\lambda_{22}}\right) - \zeta_2}{\gamma_{20}} = \frac{\left(\frac{y_3 - e_3}{\lambda_{33}}\right) - \zeta_3}{\gamma_{30}} \tag{4}$$

Como el valor esperado de los errores e_1, e_2, e_3, ζ_1, ζ_2, ζ_3 se asume nulo (ruido blanco), y se ha debido fijar la escala de cada dimensión, que suponemos es un valor unitario ($\lambda_{11} = \lambda_{22} = \lambda_{33} = 1$) entonces (5):

$$E(\eta_0) = \frac{E(y_1)}{\gamma_{10}} = \frac{E(y_2)}{\gamma_{20}} = \frac{E(y_3)}{\gamma_{30}} \tag{5}$$

Por tanto, para medir el verdadero valor[3] del factor g conviene fijar su escala, lo que ahora significa que uno de los γ debe ser también fijado. Para ello tenemos que elegir una de las dimensiones como la "mejor" o la que mejor "mide" en este caso el factor g. Si, por ejemplo, hacemos γ_{10}=1, entonces (6):

$$E(\eta_0) = E(y_1) = \frac{E(y_2)}{\gamma_{20}} = \frac{E(y_3)}{\gamma_{30}} \tag{6}$$

De este modo, con el indicador observable de la primera dimensión ya bastaría para estimar el verdadero valor del factor g[4].

Recordemos que, según Edwards (2001), este tipo de constructos con dimensiones no existen separadamente de sus dimensiones, es decir, si se admite esta conceptualización, estaríamos diciendo que el

[3] Insistimos en que no estamos necesariamente admitiendo que se pueda medir un "verdadero valor" del factor g, sino simplemente estamos empleando la misma jerga que se emplea en el uso de variables latentes en este tipo de modelización. Lo importante es lo que representan las ecuaciones relacionando las variables del modelo, más allá de que un valor numérico concreto tenga sentido o no, cuando es evidente que las escalas de medida son arbitrarias. Así, lo que se cuestiona es el modelo a través del análisis de las ecuaciones que lo representan.

[4] Si no se fijara la escala la interpretación sería idéntica, esta vez simplemente que el verdadero valor de g estaría ponderado por un coeficiente causal entre g y la dimensión pertinente, pero seguiría bastando con un único observable por cada dimensión para calcular ese valor de g.

Recibido: 14/12/2023 Aceptado: 01/06/2024

factor g no existe sin concebir su propia multidimensionalidad. Una muestra de cómo testar este tipo de modelos (incluso de tercer orden, es decir, con dimensiones y factores de orden superior en dos niveles) se encuentra en Martínez y Martínez (2008). No obstante, este tipo de modelos son abiertamente criticados por Hayduk et al. (1995), como detallan Martínez y Martínez (2010a), admitiendo las limitaciones de su trabajo anterior de 2008.

Como bien es conocido, al realizar análisis factorial para testar el modelo se está interesado en la estructura de covarianza, no en las medias. Pero ello no es óbice para admitir que, si se especifican las relaciones causales del modelo (como el presentado en la Figura 1), la interpretación sobre el valor verdadero de cada variable latente es la que hemos explicado. De este modo, proponer un modelo multidimensional como se muestra en la parte izquierda de la Figura I, que refleja un factor g subyacente, implica desde el punto de vista de la medida, el factor g puede medirse con el "mejor" indicador de la "mejor" dimensión, es decir, con un único ítem. Y más aún, debe cumplirse (6), por lo que, con cualquier ítem de las otras dimensiones, se podría conocer el valor verdadero de g tan solo con la estimación del parámetro γ respectivo.

Esto es compatible, como no podría ser de otro modo, con la visión reflectiva sobre la medición, donde los ítems de una variable latente se pueden considerar intercambiables, lo que indica que quitando un ítem no se altera el significado de la variable latente. Variaciones en la variable latente se manifiestan en variaciones en los ítems que las miden. Del mismo modo, en un constructo multidimensional, variaciones en la dimensión subyacente se manifiestan en variaciones en sus múltiples dimensiones; quitando una dimensión no se altera el significado del constructo subyacente.

Parece obvio que esta interpretación no es la que los investigadores en altas capacidades desearían plantear, pero es la que admiten cuando emplean el modelo factorial. Llegados a este punto se necesita un enfoque diferente sobre la medición, cambiando la relación causal entre los indicadores y las variables latentes, hacia modelos formativos (Bollen y Lennox, 1991; Bollen y Diamantopoulus, 2017).

Los indicadores formativos no son intercambiables, y no tienen por qué correlacionar. Si se elimina uno de esos indicadores, se altera el significado de la variable latente, ya que esa variable se define a partir de esos indicadores. De este modo, la variable latente se define a partir de los indicadores observables (con matices explicados en Bollen

y Diamantopolus, 2017), es decir, el significado de la variable latente depende de cómo se operacionalice su medición. Es una visión ciertamente constructivista, de gran utilidad práctica, que recoge perfectamente el objetivo de los modelos que plantean escalas cuyo valor total formado a partir de la agregación de ítems sirve para discriminar entre sujetos que poseen en mayor o en menor medida un determinado rasgo.

La discusión sobre los tipos de estructuras de medición formativas está fuera del objetivo de este trabajo, ya que es muy amplia su caracterización, y puede consultase en Bollen y Diamantopoulus (2017). Para el caso concreto de la escala GRS 2 analizada por Tourón et al. (2023), es quizá más adecuado remitir a propuestas como la de Hayduk et al. (2019), y que se ilustra en la parte denominada "modelo extendido" de la Figura I.

El indicador llamado $f(y_w)$ es un indicador compuesto por los ítems de las diferentes dimensiones y_1, y_2, y_3. y sus respectivos pesos w_1, w_2, w_3, es decir, es una función de los indicadores observables (7):

$$f(y_w) = \sum_{k=1}^{3} y_k w_k \tag{7}$$

Como es una combinación lineal determinista, no existe el término de error, por lo que, como indica la Figura I, $e_f = 0$.

Esta conceptualización casa mucho más con el planteamiento de Tourón et al. (2023), quienes indican:

> *"El desarrollo de la dimensión de competencia socioemocional, se llevó a cabo con el objetivo de ampliar la evaluación de la alta capacidad, más allá de una lente tradicional que se enfoca principalmente en las "fortalezas de la cabeza" -que incluyen resolución de problemas, memoria, razonamiento y creatividad-, a una visión más holística y completa del estudiante que incluye las "fortalezas del corazón", como las fortalezas personales e interpersonales (Pfeiffer, 2001, 2017b). Esencialmente, el propósito era incorporar una perspectiva de psicología positiva a la escala para padres GRS 2".*

Aquí son los investigadores los que definen qué es alta capacidad a partir de la instrumentalización de la medida. Es decir, y expresado de manera sencilla, se incluye una nueva dimensión porque creen que al hacerlo se define o evalúa mejor el concepto de alta capacidad. Por tanto, la evaluación de la alta capacidad dependerá de las dimensiones elegidas para su medición, y esta es una visión formativa de la medición,

no reflectiva. De este modo, la caracterización de alta capacidad en un individuo dependerá de las puntuaciones en las diferentes dimensiones. Así, aquellos que obtengan puntuaciones altas en, por ejemplo, todas las dimensiones serán considerados como que tienen una mayor capacidad que aquellos que obtengan puntuaciones altas en una sola dimensión. En consecuencia, el valor de esa variable tan elusiva que se quiere medir como capacidad general depende de las puntuaciones en todas las dimensiones; se realiza un perfil de la capacidad de cada persona en función de su puntuación en cada dimensión y en la combinación de estas.

La última parte del modelo extendido de la Figura I la componen las variables latentes η_4 y η_5, que corresponden a efectos de la escala de medida y a una variable de control, respectivamente. Es decir, se expande el modelo factorial al considerar una estructura causal entre variables latentes, donde se tengan en cuenta los efectos de la escala de medida de altas capacidades y una (o varias, dado el caso) covariables exógenas que actúen como control de esos efectos. Este esquema, es similar al postulado por Hayduk et al. (2019), por lo que se puede emplear el enfoque descrito en ese trabajo para testar el modelo GRS 2.

Analogía con las altas capacidades físicas

El modelo presentado en la Figura I que representa un instrumento de medida de las altas capacidades intelectuales, es quizá más entendible si se postula su equivalente a la medición de las altas capacidades físicas.

Partamos de que η_0 es una variable que representa la alta capacidad física que es un concepto elusivo y difícil de medir directamente. Entonces, se plantea una concepción multidimensional, en el que otras variables latentes entran en juego: por ejemplo, la velocidad, la fuerza y la resistencia aeróbica (Jung, 2022). Esas variables se miden a partir de indicadores observables (test de velocidad, test de fuerza, test de resistencia). Y esto es no es una conceptualización reflectiva, sino formativa, por las siguientes razones:

- Estamos definiendo la alta capacidad física a partir de esas 3 dimensiones que consideramos. Otro investigador podría desafiar este planteamiento y añadir otra dimensión, por ejemplo, la flexibilidad (ver Jung, 2022) con lo que la definición de alta capacidad física cambiaría.

Recibido: 14/12/2023 Aceptado: 01/06/2024

- Las dimensiones consideradas no tienen por qué estar fuertemente asociadas, ya que aquellos que tengan más masa muscular y más fibras rápidas probablemente sean mejores en velocidad y fuerza, pero no tanto en resistencia.
- Las dimensiones no son intercambiables, y si se elimina una de ellas se altera el significado de la variable de interés (la alta capacidad física).
- La alta capacidad física no se puede medir a través de un indicador, como se desprende de la concepción reflectiva de la ecuación (6). Es decir, sería incorrecto estimar la alta capacidad física con solo la medición de una de las dimensiones, por lo que necesitamos todas las dimensiones sobre las que se ha definido el concepto de alta capacidad física para "medirlo".
- Aquellos individuos que puntúen más alto en las 3 dimensiones serán considerados que tienen mayor capacidad física que los que puntúen alto en una dimensión, pero bajo en las otras dos.

De este modo, se podría construir una escala con una función de pesos a especificar por el investigador para formar el índice global $f(y_w)$. Y este índice se podría relacionar causalmente con variables de consecuencias η_5, como, por ejemplo, el rendimiento del deportista en un equipo de fútbol, y añadir variables de control η_4, como, por ejemplo, el sexo o la edad.

Estos argumentos, aunque simplistas, son didácticos, y nos hacen ver que la medición de las altas capacidades intelectuales debe seguir una modelización similar.

Test del modelo

Tourón et al. (2023) reportan la bondad de ajuste de su modelo original, y de las posteriores especificaciones que testan (comparan el ajuste de 8 modelos competitivos), no empleando el test de la chi-cuadrado como criterio[5] (ajuste exacto) y, en su lugar, usando otra serie de índices de ajuste, entre ellos el RMSEA, TLI y otros: *"Siguiendo a Hu y Bentler*

[5] Recordemos que el valor de la chi-cuadrado dividido por los grados de libertad, que sí reportan los autores, no es ningún test estadístico.

Recibido: 14/12/2023 Aceptado: 01/06/2024

(1999), como evidencia de validez es suficiente un ajuste aceptable en la combinación de estos índices".

Pocos años más tarde del estudio de Hu y Bentler (1999), diversos trabajos (Beaducel y Witttman, 2005; Fan y Sivo, 2005; Marsh et al., 2004) cuestionaron el uso de criterios fijos, o "reglas de oro" para la evaluación de modelos con índices de ajuste "aproximado" (CFI, TLI, SRMR, RMSEA, etc.) como respuesta a las reglas de corte propuestas por Hu y Bentler (1999). Otros estudios posteriores también lo hicieron (ej. Niemand y Mai, 2018; McNeish y Wolf, 2023). Esos criterios comúnmente son usados para ajustar los modelos de forma que se evite el único test adecuado para detectar la mala especificación de un modelo (el test de la chi-cuadrado o, más correctamente, la familia de test de la chi-cuadrado) prueba que nos dice que, si el modelo es correcto, no será rechazada incluso con tamaños de muestra más grandes (ej. Hayduk y Glaser, 2000a,b; Barret, 2007; Hayduk et al., 2007; McIntosh, 2007; Antonakis et al., 2010; McIntosh 2012; Hayduk, 2014b; Ropovic, 2015; Rönkkö et al., 2016).

El grado de mal ajuste en la estructura de covarianza no está asociado al grado de mala especificación (Hayduk, 2014b), de este modo es incorrecto identificar la divergencia estadística sobre el modelo propuesto como un indicador de grado de idoneidad de este. En los modelos de ecuaciones estructurales se parte de una matriz de covarianzas entre los indicadores observables. Esa matriz se compara con la matriz de covarianzas implicadas por el modelo, la cual resulta de las restricciones de causalidad que estemos imponiendo. Esa matriz implicada por el modelo es la matriz poblacional, porque estipulamos que ése es el modelo que rige la población. Por tanto, el modelo especificado es la visión del mundo que queremos testar contra los datos, los cuales, salvo la propia variabilidad muestral (de ahí el test estadístico de la chi-cuadrado en función de sus grados de libertad), deben ajustarse exactamente.

El test de la chi-cuadrado puede fallar en ocasiones en detectar modelos mal especificados. Por ejemplo, es bien conocida la posible existencia de modelos equivalentes (mismo ajuste, pero con diferentes restricciones causales). De ahí que el compromiso teórico con el modelo sea tan importante para elegir entre modelos que se ajustan aquel que es congruente con la teoría que lo soporta. No obstante, y pese a esa limitación, sigue siendo la única vía posible de análisis, mejor que el resto de los índices que comúnmente se reportan, casi siempre citando el trabajo de Hu y Bentler (1999), como bien explica Hayduk (2014a).

De este modo, cuando un modelo no pasa el test de la chi-cuadrado, los parámetros estimados no son interpretables, ya que una o varias de las relaciones de covarianza implicadas por el modelo no es apoyada por los datos empíricos, lo que de inmediato produce sesgo en las estimaciones. Obviamente, no tiene sentido tampoco interpretar índices como la fiabilidad compuesta o la varianza media extraída. Si, por el contrario, el modelo se ajusta, entonces hay que considerar la posibilidad de modelos competitivos (teorías alternativas) y/o equivalentes que se puedan ajustar también, y elegir el que tenga un mejor ajuste en el caso de modelos en competencia (vía test de la diferencia de chi-cuadrado entre dos modelos competitivos que se ajustan) o el que sea teóricamente fundado en el caso de los modelos equivalentes (en línea con la concepción de causalidad que parte de aseveraciones cualitativas estipulada por Pearl, 2000).

Como se deriva de las anteriores líneas, este enfoque es mucho más exigente con los modelos que se plantean, sobre todo en ciencias sociales, donde existe una especie de "escritura automática" en investigación (Martínez y Martínez, 2009); se van reproduciendo errores e incorrecciones sistemáticamente, lo que hace que esos errores sean la postura mayoritaria, y esa postura, simplemente por ser mayoritaria, siga incitando a que otros la adopten. Hayduk (2014b) ilustra este hecho con una anécdota en el que uno de los creadores del software que fue, durante muchos, años el más empleado en esta metodología (LISREL) admitió que, como el test de la chi-cuadrado rechazaba muchos modelos, tuvieron que proponer otros indicadores de ajuste (GFI, AGFI, etc.) menos exigentes con los datos para que hubiera menos modelos rechazados, contentar a los investigadores y, de este modo, que el software tuviera más aceptación.

Análisis factorial

Tourón et al. (2023) emplean el análisis factorial exploratorio y confirmatorio: *"Para probar la estructura de la escala se ha llevado a cabo un análisis factorial exploratorio y otro confirmatorio que aporten evidencias sobre la validez de la misma"*.

Hayduk (2014b) demuestra que el análisis factorial exploratorio es incapaz de detectar la estructura real de los datos, es decir, de identificar

Recibido: 14/12/2023 Aceptado: 01/06/2024

el modelo que ha generado esos datos empíricos. Por tanto, no se puede *probar* un modelo teórico con un análisis exploratorio, sino simplemente tener una visión de la estructura correlacional de los datos.

Tourón et al. (2023) hacen un meritorio esfuerzo en dividir la muestra en dos partes. Sobre el primer 40% de la muestra emplean el análisis factorial exploratorio y sobre el 60% restante testan la estructura de 4 factores que encuentran en la fase exploratorio. Desde el punto de vista metodológico es loable esa filosofía de particionar los datos, aunque hubiera sido más correcto emplear las mismas particiones, pero de forma diferente. Así en la primera se podría haber puesto a prueba la estructura original de 3 factores mediante un análisis factorial confirmatorio, estudiando las posibles causas de desajuste (analizando, por ejemplo, los índices de modificación siempre que estén sustentados en la teoría). Si, tras analizar los índices de modificación, se hubiera recomendado una especificación diferente que se ajustara vía chi-cuadrado, entonces en el 60% restante de la muestra se podría haber testado esa nueva especificación con un nuevo análisis factorial confirmatorio.

Esta propuesta que acabamos de hacer dista, sin embargo, de ser la más óptima (aunque es mejor que realizar un análisis factorial exploratorio, ya que es más congruente con cualquier método hipotético-deductivo que pretenda *probar* la estructura de una escala). Aunque el test de la chi-cuadrado es más efectivo que la identificación de valores propios en el análisis factorial exploratorio (Hayduk, 2014b), también puede, en ocasiones, fallar al identificar el modelo correcto. Es entonces cuando, indefectiblemente, surge la necesidad de analizar la validez de constructo a través de la construcción de una red de causas y/o efectos de las variables latentes sobre las que se esté estudiando su validez, es decir, postular una red nomológica, tal y como Cronbach y Meel (1955) también defendían.

Como vuelve a recordar Hayduk (2014b) poniendo a prueba el constructo en esa red nomológica se interroga la capacidad del factor latente de producir sus indicadores observables, y eso es indispensable para hablar de validez de constructo (Cronbach y Meehl, 1955). De este modo, en todo estudio en el que se pretenda validar un instrumento de medida, debe aparecer un "modelo extendido" según la terminología que hemos empleado en la Figura I, es decir, ir más allá del análisis factorial e incluir el constructo o los constructos de interés en una red nomológica. Es entonces cuando los modelos de ecuaciones estructurales llegan a

su máximo potencial, pudiendo poner a prueba ese modelo contra los datos empíricos, es decir, la matriz de covarianzas implicada por las restricciones causales del modelo contra la matriz de covarianza de los indicadores observables.

Existen numerosas críticas a una forma de proceder en psicometría que, no por extendida, es correcta. Se trata de tratar de validar un instrumento de medida con análisis factorial exploratorio (aunque luego se haga un análisis factorial confirmatorio), o con un análisis factorial confirmatorio (aunque luego se teste el modelo causal, lo que es conocido como "análisis en dos pasos"). La razón fundamental de esas críticas es la relativa al concepto de validez de constructo, explicado en Cronbach y Meel (1955); la medición no se puede separar de la teoría, y la validez de constructo necesita de una red nomológica para poder ser analizada. Ejemplos de esas críticas se pueden encontrar en Fornell y Yi (1992), Hayduk y Glasser (2000a) o Hayduk (2014b).

Número de ítems por variable latente

Tourón et al. (2023) tradujeron los 20 ítems originales de la escala GRS 2 para padres, agrupados en las 3 dimensiones mencionadas (6, 7, y 7, respectivamente). Como Tourón et al. (2023) no indican la redacción de los ítems es muy complejo valorar la idoneidad de los mismos para reflejar la dimensión subyacente. No obstante, y a estas alturas de este manuscrito, ya sabemos que en un modelo reflectivo el valor verdadero de cada dimensión latente se puede estimar a partir del valor de cada indicador observable (1). Los indicadores observables son reflectivos según la especificación de Tourón et al. (2023) por lo que con solo un indicador se podría medir cada una de esas dimensiones. Si un investigador piensa que con un solo indicador es insuficiente para medir una variable latente y que necesita más porque la dimensión latente es *amplia* (como indican Tourón et al. (2023)), lo que probablemente está sucediendo es que esa batería de indicadores está midiendo variables latentes diferentes, es decir, no son la manifestación de una única variable latente, sino de varias.

El usar un solo indicador por variable latente en la Figura I no es solo un modo de simplificar el esquema, sino que es una forma de reflejar las investigaciones que defienden las medidas de un solo ítem por variable

Recibido: 14/12/2023 Aceptado: 01/06/2024

latente cuando, evidentemente, la variable latente se manifiesta a través de observables reflectivos (ej. Bergkvist y Rossiter, 2007; Matthews et al., 2022; Allen y Pistone, 2023; Wulf et al., 2023). Como indican Durvasula et al. (2012), hay literatura que sugiere que las escalas largas amenazan a la validez de constructo, y producen fatiga en la persona que responde. Además, cuando se emplea el mismo formato de escala Likert, es decir, el mismo rango a una numerosa batería de ítems, existe riesgo serio de sesgo de método común (Jordan y Troth, 2020), y una de sus manifestaciones puede ser precisamente lo que Tourón et al. (2023) parecen considerar como algo positivo cuando aluden a las correlaciones altas entre cada ítem y el resto señalando la *"homogeniedad del conjunto de ítems"*. Es más, hay ejemplos en la propia literatura sobre altas capacidades y escalas GRS donde se vislumbra un efecto halo que afecta obviamente a la validez (ej. Jabůrek et al., 2020).

Una explicación exhaustiva de la necesidad de emplear un solo ítem (el mejor indicador posible), o en determinados casos, un segundo ítem que acompañen a ese mejor indicador puede encontrarse en Hayduk (1996) o en Hayduk y Littvay (2012). En el contexto de ecuaciones estructurales, Hayduk (1996) propone además fijar la varianza de error de ese mejor indicador, haciendo que el investigador se comprometa con el significado de la variable latente, evitando de esta forma el denominado *interpretational confounding*, es decir, el cambio entre la relación entre la variable latente y su mejor indicador cada vez que cambia la estructura de covarianza de los datos al añadir otras variables en el modelo.

De este modo, en la Figura I sería necesario no solo fijar todos los λ (que hemos dicho que era primordial para escalar cada variable latente), sino también las varianzas de error de cada indicador, usualmente a un porcentaje muy reducido de lo que sería la varianza del ítem observable. El modelo representado en la Figura I se podría ampliar a dos ítems por variable latente. De hecho, esto añadiría robustez a la conceptualización porque ese segundo ítem (el segundo mejor indicador posible) no tendría fijada ni la carga factorial ni la varianza de error, por lo que las nuevas covarianzas añadidas a la matriz de covarianzas muestral producirían un test más sólido de las restricciones producidas por la fijación de la carga factorial y de la varianza de error de esos mejores indicadores.

Como explican Hayduk y Littvay (2012) si los indicadores son reflectivos no es conveniente utilizar tantos ítems por variable latente, porque lo que se hace es saturar la parte latente del modelo con

correlaciones factoriales. Si se necesitan un gran número de ítems por variable latente porque desde el punto de vista teórico el dominio de la variable es amplio, insistimos en que lo más probable es que, o se estén midiendo en realidad varias variables latentes, o se esté vislumbrando una estructura formativa en lugar de reflectiva.

Modelos alternativos

Existen otras especificaciones de modelos compatibles con toda la argumentación anterior, y que son diferentes a la Figura I, tal y como se muestran en la Figura II, III y IV.

En la Figura II se representa un modelo similar al de la Figura I, salvo que ahora no se pretende construir ningún índice global que mida la capacidad como función de las dimensiones, sino que la teoría es más prudente y admite que existen dimensiones η_1, η_2, η_3 que se pueden medir de forma reflectiva con y_1, y_2, y_3, asociadas a las altas capacidades, aunque no se especifica cuál es ese tipo de relación funcional (de ahí que es una especificación más conservadora). Esas dimensiones influyen en una variable consecuencial η_5 (por ejemplo,

FIGURA II. Modelo alternativo sin escala global

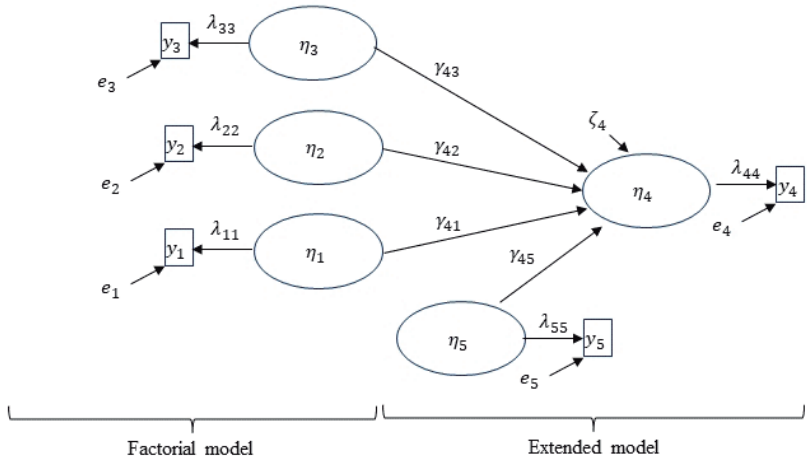


FIGURA III. Modelo alternativo sin escala global y con causas de dimensiones

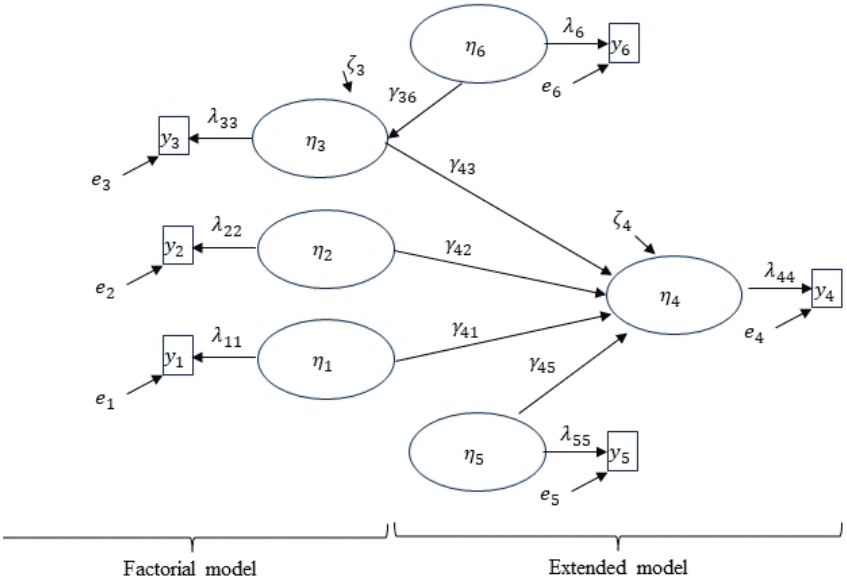
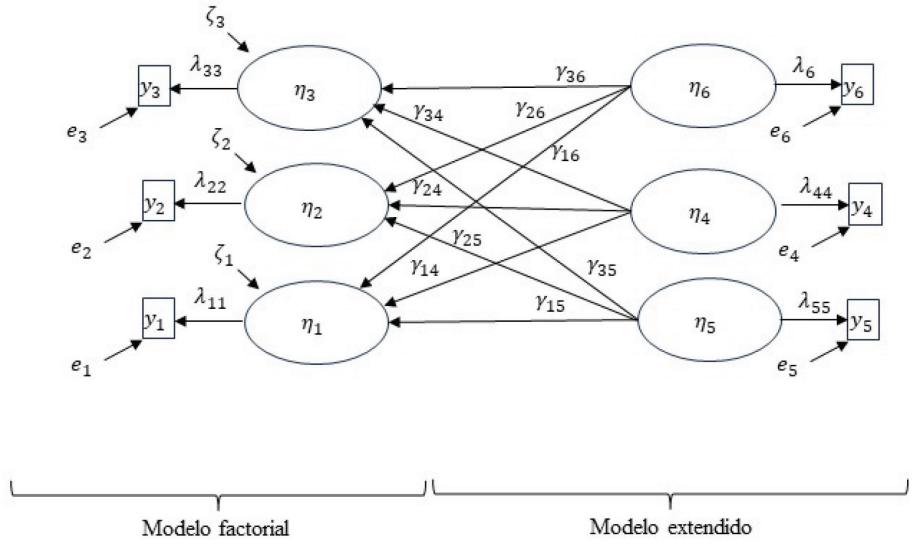


FIGURA IV. Modelo alternativo con causas de las dimensiones de altas capacidades



Recibido: 14/12/2023 Aceptado: 01/06/2024

alguna variable de rendimiento, como especifican Kelly y Peverly, 1992) de la que se puede medir su efecto a través de los coeficientes γ, y en la que hay una variable de control η_4 (o varias, aunque por simplicidad solo se ha propuesta una) que podría ser, por ejemplo, alguna variable socioeconómica o demográfica.

En la Figura III se amplía aún más el modelo propuesto en la Figura II, siendo todavía más interesante desde el punto de vista causal porque incluye una variable η_6 que puede provocar variación en una de las dimensiones, como por ejemplo el contexto familiar sobre la creatividad (Manzano y Arranz, 2008).

En la Figura IV no se muestran consecuencias de las dimensiones de altas capacidades, sino posibles causas. Este esquema es (con matices) equivalente al utilizado por Pfeiffer et al. (2008) para validar la escala GRS-S, proponiendo variables que pueden causar variación en las dimensiones que ellos emplean (6 en lugar de las 3 de la Figura IV), donde como esas variables latentes exógenas η_4, η_5, η_6, emplean el género, la raza y la edad. Evidentemente lo hacen usando otra metodología estadística (análisis multigrupo de la varianza), pero el diagrama causal es equivalente al mostrado en la Figura IV. Para testar la hipótesis que Pfeiffer et al. (2008) estipulan de que la escala GRS-S no debería discriminar entre género, raza y edad, en un modelo de ecuaciones estructurales simplemente habría que fijar todos los parámetros γ a cero, y poner a prueba esas restricciones de causalidad con el test de la chi-cuadrado. Recordemos que los modelos de ecuaciones estructurales permiten emplear variables exógenes dicotómicas, y que cualquier variable es susceptible de ser considerada latente debido al error de medición. Por ejemplo, el sexo o la edad pueden tener errores de codificación, por lo que se puede fijar la varianza de error de sus indicadores no necesariamente a cero.

Hay que matizar que Pfeiffer et al. (2008) no emplean un ítem para medir cada dimensión, sino varios (12 por cada una de las 6 dimensiones), con lo que las críticas realizadas en este manuscrito con relación a la confusión entre indicadores reflectivos y formativos, y lo que ello implica para la modelización estarían vigentes.

Por tanto, cualquiera de las propuestas de modelización expresadas en las Figuras II, III y IV, hubieran ayudado a Tourón et al (2023) a analizar más rigurosamente la validez de constructo de su adaptación de la escala GRS 2 al entorno español.

Recibido: 14/12/2023 Aceptado: 01/06/2024

Discusión

El estudio de Tourón et al. (2023), aunque realiza un esfuerzo procedimental meritorio para analizar la validez de constructo de la escala GRS 2 para padres en el ámbito español, adolece de elementos esenciales que cuestionan su validez.

El primero de ellos se refiere a la especificación de un modelo reflectivo en lugar de formativo. Tourón et al. (2023) consideran tanto los ítems de cada dimensión como las propias dimensiones de carácter reflectivo (Edwards, 2001). Así los especifican en la relación causal entre las variables latentes y sus respectivos ítems, y en la relación entre los factores de segundo orden y las dimensiones. Aunque los autores no muestran en su artículo la batería de ítems completa, sí que comentan alguno de ellos, como los ítems 5, 8 y 15 que los autores sitúan como manifestación de la dimensión (al final identificada como subdimensión) "control emocional". Esos ítems se refieren al control del estrés, la ira y la perseverancia en el trabajo, lo que precisamente sugiere que son indicadores de variables latentes diferentes o que deberían ser causa y no efecto de esa variable etiquetada (construida y definida en base a sus indicadores) "control emocional". El construir escalas sumativas con los ítems de cada dimensión o en una única escala global para dar un perfil de capacidad de cada individuo es una visión formativa, cuya perspectiva de análisis se muestra en la Figura I, requiriendo, por tanto, añadir a la estructura factorial otras variables que se relacionen con la escala o escalas construidas. Esta visión formativa estaría en consonancia con la visión de Boring (1923) sobre la inteligencia y los test para medirla[6].

El segundo aspecto problemático del estudio de Tourón et al. (2023) es la no admisión de que todos los modelos testados (ocho en total) son incompatibles con los datos empíricos y, por tanto, los modelos son incorrectos, es decir, no válidos. Es cierto que en la literatura especializada de ecuaciones estructurales hay investigadores (ej. Jöreskog, 1978; Bentler, 1990; Steiger; 2007; Browne et al., 2002) que critican la prueba chi-cuadrado y recomiendan los índices de ajuste

[6] De nuevo no pretendemos entablar una discusión más profunda sobre los conceptos de inteligencia, capacidad, etc. Solo argumentamos que una visión formativa nos diría que lo que estamos midiendo viene determinado por cómo lo hemos definido previamente. Por tanto, la clave está en cómo se defina el concepto que se quiere medir.

aproximados. Curiosamente Jöreskog, Bentler, Steiger y Browne fueron creadores de software comercial de ecuaciones estructurales (LISREL, EQS, SEPATH, y RAMONA respectivamente). Otros investigadores siguen en esa misma línea (ej. McNeish y Wolf, 2023). Sin embargo, las principales críticas al test de ajuste exacto son respondidas una a una por Hayduk (2014b). Una de las más recurrentes críticas es el mantra de que "todos los modelos son falsos", por lo que no tendría sentido buscar el ajuste exacto. Aunque Hayduk (2014b) responde también a esa crítica, conviene consultar fuera del ámbito de las ecuaciones estructurales, y en un campo más general sobre teoría de probabilidad y modelización de datos e inferencia estadística (ej. Spanos, 2019). Precisamente Spanos (2021) recuerda que existe una mala interpretación de ese "mantra" ya que hay que saber distinguir entre un modelo sustantivo (basado en una teoría) y el modelo estadístico implícito en él, en el cual se encuadran un conjunto de asunciones probabilísticas impuestas sobre los datos. Así, los investigadores que aluden a que "todos los modelos son falsos" intentan justificar que la mala especificación estadística es inevitable y, por tanto, no tiene sentido buscar el ajuste exacto. Pero eso es erróneo, porque ese lema se refiere a los modelos sustantivos que, obviamente, son una representación más simplificada de la realidad. Sin embargo, lo que se testa vía chi-cuadrado en los modelos de ecuaciones estructurales no es si el modelo sustantivo es "verdad", sino si el modelo estadístico implícito es congruente con los datos empíricos, el cual está compuesto inseparablemente de las restricciones causales y de las asunciones probabilísticas. De ahí la importancia de la teoría en construir un modelo causal, y de ahí que haya que ser rigurosos en respetar la evidencia marcada por la discrepancia entre el modelo propuesto y los datos empíricos cuando las asunciones del modelo se cumplen, tal y como Spanos (2019) explica de forma general en la modelización de datos.

Esto nos lleva indefectiblemente al tercer elemento de crítica a Tourón et al. (2023) relacionado con el uso de análisis factorial para el estudio de la validez de constructo, y la necesidad de estipular una red nomológica que pueda ser testada globalmente. Bien es cierto que, como se indica en Lissitz (2009), existen también propuestas sobre el abandono de los procesos de validación recomendados por Cronbach y Meehl (1955) y enfocarse en una perspectiva ontológica, como la defendida en Borsboom et al. (2004): un test es válido para medir un atributo si y sólo si (a) el atributo existe y (b) variaciones en el atributo causalmente producen

Recibido: 14/12/2023 Aceptado: 01/06/2024

variaciones en los resultados del procedimiento de medición. Sin embargo, las tesis de Hayduk (1996) y Borsboom et al. (2004) comparten más de lo que disienten, como analizan Martínez y Martínez (2009), por lo que no debe haber disconformidades sustanciales con los argumentos expuestos en este trabajo.

El número de ítems por variable latente es la cuarta crítica a Tourón et al. (2023). Aunque, como se ha indicado anteriormente, es más que probable que varios de los ítems del instrumento GRS 2 sean formativos, si se estipulan como reflectivos no hay necesidad de saturar el instrumento con tantos ítems que puedan provocar diferentes tipos de sesgo. Sin embargo, si los investigadores defienden la necesidad de todos esos ítems por la amplitud de los conceptos medidos y porque son necesarios para recoger las características de lo que se quiere medir entonces implícitamente se estaría admitiendo una estructura formativa, no reflectiva.

Hemos propuesto, además, modelos diversos que siguen la misma filosofía que muestra la Figura I, y que se cubrirían las limitaciones de la propuesta de Tourón et al. (2023). No son únicos, ya que existen otras propuestas en investigación en altas capacidades que ya se aventuran por especificaciones cercanas, como Nakano et al. (2016) y su modelo MIMIC (múltiples causas y múltiples indicadores). Es más, como también hemos indicado Pfeiffer et al. (2008) emplea una forma similar (aunque no tan rigurosa estadísticamente) de análisis de validación de la escala GRS-S como la especificada en la Figura IV.

En esos modelos de las Figuras II, III y IV no se plantea explícitamente que haya un factor de orden superior de capacidad que se manifieste a través de dimensiones. Una futura línea de investigación atrayente sería realizar un enfoque similar al de redes en psicopatología (Borsboom y Cramer, 2013; Fonseca-Pedrero, 2017), considerando esas dimensiones de capacidad como "síntomas" que se relacionan complejamente en una red de interacción. Ninguna de las concepciones que hemos propuesto es incompatible con esa visión, ya que se dejan libres las correlaciones entre esas dimensiones de capacidad.

Los ejemplos de modelización mostrados en las Figuras I a IV son solo una referencia. Obviamente deben adaptarse al conjunto de variables latentes de cada contexto. Sin embargo, conviene aclarar que se necesita especificar muy bien qué se está midiendo y qué variable se consideran como un tipo de dotación o no. Tourón et al. (2023) indican que las

habilidades socioemocionales no se ven como un tipo de dotación, por tanto, a la hora de explicar qué significa multidimensionalidad en altas capacidades y especificarlo en un modelo hay que tener clara esa distinción.

Otras críticas al instrumento de medida GRS 2 también serían importantes de debatir, aunque ya superarían la extensión de este trabajo. Entre ellas, por ejemplo, estaría la discusión sobre el uso de escalas tipo Likert ordinales con etiquetas lingüísticas, dado que puede haber interacción entre esas etiquetas lingüísticas y la numeración de la escala Likert. A este respecto, enfoques basados en lógica borrosa podrían ayudar a dilucidar la idoneidad de esas escalas (ver Martínez y Martínez, 2010b; Martínez et al., 2010). En cualquier caso, reportar valores medios, como hacen Tourón et al (2023) en escalas ordinales es, cuanto menos, cuestionable (Liddell y Kruschke, 2018).

Finalmente, y dado que Tourón et al. (2023) admiten que su estudio es una validación preliminar y que se debe corroborar con investigaciones posteriores, sugerimos que se consideren los postulados argumentados en este manuscrito para seguir avanzando en la medición de las altas capacidades.

Referencias bibliográficas

Allen, B., & Pistone, L. F. (2023). Psychometric evaluation of a single-item screening tool for the presence of problematic sexual behavior among preteen children. *Child abuse & neglect, 143,* 106327. https://doi.org/10.1016/j.chiabu.2023.106327

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R. (2010). On making causal claims: A review and recommendations. *Leadership Quarterly, 21,* 1086-1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Barret P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*(5), 815-824. https://doi.org/10.1016/j.paid.2006.09.018

Beauducel, A., & Wittmann, W. W. (2005). Simulation Study on Fit Indexes in CFA Based on Data With Slightly Distorted Simple Structure. *Structural Equation Modeling, 12*(1), 41–75. https://doi.org/10.1207/s15328007sem1201_3

Recibido: 14/12/2023 Aceptado: 01/06/2024

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bergkvist, L. & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. Journal of Marketing Research, 44 (2), 175-184. https://doi.org/10.1509/jmkr.44.2.175

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), 305–314. https://doi.org/10.1037/0033-2909.110.2.305

Bollen K. A. (2002). Latent variables in psychology and the social sciences. *Annual review of psychology, 53,* 605–634. https://doi.org/10.1146/annurev.psych.53.100901.135239

Bollen, K. A., & Diamantopoulos, A. (2017). In defense of causal-formative indicators: A minority report. *Psychological Methods, 22*(3), 581–596. https://doi.org/10.1037/met0000056

Boring, E. G. (1923). Intelligence as the test measures it. *The New Republic, 35,* 35-37.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9,* 91–121.

Borsboom, D., Mellenbergh, G. J., van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*(4), 1061-1071.

Browne, M. W., MacCallum, R. C., Kim, C. T., Andersen, B. L., Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods, 7*(4), 403–421. https://doi.org/10.1037//1082-989x.7.4.403

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological test. *Pyschological Bulletin, 52,* 218-302. https://doi.org/10.1037/h0040957

Durvasula, S., Sharma, S., Carter, K. (2012). Correcting the t statistic for measurement error. *Marketing Letters, 23*(3), 671–682. https://doi.org/10.1007/s11002-012-9170-9

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: Towards an integrative and analytical framework. *Organizational Research Methods, 4,* 144–192.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index

strategy revisited. *Structural Equation Modeling, 12,* 334-367. https://doi.org/10.1207/s15328007sem1203_1

Fonseca-Pedrero, E. (2017). Análisis de redes: ¿una nueva forma de comprender la psicopatología? *Revista de Psiquiatría y Salud Mental, 10*(4), 206-215.

Fornell, C. & Yi, Y. (1992). Assumptions of the Two-Step Approach to Latent Variable Modeling. *Sociological Methods & Research, 20,* 291-320.

Hayduk, L. A. (1996). *LISREL Issues, Debates and Strategies*. Baltimore, MD: Johns Hopkins University Press.

Hayduk, L. A. (2014a). Seeing perfectly-fitting factor models that are causally misspecified: Understanding that close-fitting models can be worse. *Educational and Psychological Measurement, 74*(6), 905-926. https://doi.org/10.1177/0013164414527449

Hayduk, L. A. (2014b). Shame for disrespecting evidence: The personal consequences of insufficient respect for structural equation model testing. *BMC: Medical Research Methodology, 14,* 124. https://doi.org/10.1186/1471-2288-14-124

Hayduk, L. A., Cummings, G. Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three – Testing the theory in structural equation models! *Personality and Individual Differences, 42*(5), 841-850.

Hayduk, L. A., Estabrooks, C. A., Hoben, M. (2019). Fusion Validity: Theory-Based Scale Assessment via Causal Structural Equation Modeling. *Frontiers in Psychology, 10,* 1139. https://doi.org/10.3389/fpsyg.2019.01139

Hayduk, L. A., & Glaser, D. N. (2000a). Jiving the four-step, waltzing around factor analysis, and other serious fun. *Structural Equation Modeling, 7*(1), 1-35.

Hayduk, L. A., & Glaser, D. N. (2000b). Doing the four-step, right-2-3, wrong-2-3: A brief reply to Mulaik and Millsap; Bollen; Bentler; and Herting and Costner. *Structural Equation Modeling, 7*(1), 111-123.

Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? BMC Medical *Research Methodology, 12,* 159, 1-17. https://doi.org/10.1186/1471-2288-12-159

Hayduk, L.A., Ratner, P.A., Johnson, J.L., Bottorff, J.L. (1995). Attitudes, ideology and the factor model. *Political Psychology, 16,* 479–507.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jabůrek, M., Ťápal, A., Portešová, Š., Pfeiffer, S. I. (2020). Validity and Reliability of Gifted Rating Scales-School Form in Sample of Teachers and Parents – A Czech Contribution. *Journal of Psychoeducational Assessment, 39*(3), 361–371. https://doi.org/10.1177/0734282920970718

Jordan, P. J., & Troth, A. C. (2020). Common method bias in applied settings: The dilemma of researching in organizations. *Australian Journal of Management, 45*(1), 3-14. https://doi.org/10.1177/0312896219871976

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika, 43,* 443-477.

Jung J. Y. (2022). Physical giftedness/talent: A systematic review of the literature on identification and development. *Frontiers in Psychology, 13,* 961624. https://doi.org/10.3389/fpsyg.2022.961624

Kelly, M. S. & Peverly, S. T. (1992). Identifying bright kindergartners at risk for learning difficulties: Predictive validity of a kindergarten screening tool. *Journal of School Psychology, 30,* 245-258.

Liddell, T. M. & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology, 79,* 328-348. https://doi.org/10.1016/j.jesp.2018.08.009

Lissitz, R. W. (Editor) (2009). *The Concept of Validity: Revisions, New Directions, and Applications*. Charlotte, NC: Information Age Publishing.

MacCallum, R. C., & Austin, J. T. (2000). Applications of Structural Equation Modeling in Psychological Research. *Annual Review of Psychology, 51,* 201-226. https://doi.org/10.1146/annurev.psych.51.1.201

Manzano, A. Arranz, F., Enrique B. (2008). Contexto familiar, superdotación, talento y altas capacidades». *Anuario de psicología / The UB Journal of psychology, 39*(3), 289-309, https://raco.cat/index.php/AnuarioPsicologia/article/view/123643

Marsh, H. W., Hau, K. T., Wen, Z. (2004). In search of golden rules: Comment of hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320-341. doi:10.1207/s15328007sem11032

Martínez, J. A., Ko, Y. J., & Martínez, L. (2010). An application of fuzzy logic to service quality research: A case of fitness service. *Journal of Sport Management, 24,* 502-523. https://doi.org/10.1123/jsm.24.5.502

Martínez, J. A. & Martínez, L. (2009). El análisis factorial confirmatorio y la validez de escalas en modelos causales. *Anales de Psicología, 25*(2), 368-374.

Martínez, J. A. & Martínez, L. (2010a). Re-thinking perceived service quality. An alternative to hierarchical and multidimensional models. *Total Quality Management & Business Excellence*, 21(1), 93-118

Martínez, J. A. & Martínez, L. (2010b). La medición de la satisfacción de servicios deportivos a través de la lógica borrosa. *Revista de Psicología del Deporte, 19(1)*, 41-58.

Martínez, L., & Martínez, J. A. (2008). Developing a multidimensional and hierarchical service quality model for the travel agencies industry. *Tourism Management, 29,* 706–720.

Matthews, R. A., Pineault, L. & Hong, Y. H. (2022). Normalizing the use of single-item measures: Validation of the single item compendium for organizational psychology. *Journal of Business and Psychology, 37,* 639-673. https://doi.org/10.1007/s10869-022-09813-3

McIntosh, C. N. (2007). Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007). *Personality and Individual Differences, 42*(5), 859–867. https://doi.org/10.1016/j.paid.2006.09.020

McIntosh, C. N. (2012). Improving the evaluation of model fit in confirmatory factor analysis: A commentary on Gundy, CM, Fayers, PM, Groenvold, M., Petersen, M. Aa., Scott, NW, Sprangers, MAJ, Velikov, G., Aaronson, NK (2011). Comparing higher-order models for the EORTC QLQ-C30. *Quality of Life Research. Quality of Life Research, 21*(9), 1619- 1621.

McNeish, D., & Wolf, M. G. (2023). Dynamic fit index cutoffs for confirmatory factor analysis models. *Psychological Methods, 28*(1), 61–88. https://doi.org/10.1037/met0000425

Nakano, T. D., Primi, R., Ribeiro, W. D., Almeida, L. D. (2016). Multidimensional Assessment of Giftedness: criterion Validity of Battery of Intelligence and Creativity Measures in Predicting Arts and Academic Talents. *Anales De Psicologia, 32,* 628-637.

Niemand, T. & Mai, R. (2018). Flexible Cutoff Values for Fit Indices in the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science, 46*(6), 1148-1172.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Pfeiffer, S. I., Petscher, Y., & Kumtepe, A. (2008). The Gifted Rating Scales-School Form: A Validation Study Based on Age, Gender, and Race. *Roeper review, 30*(2), 140–146. https://doi.org/10.1080/02783190801955418

Rönkkö, M., McIntosh, C. N., Antonakis, J., Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management, 47–48,* 9–27. https://doi.org/10.1016/j.jom.2016.05.002

Ropovik, I. (2015). A cautionary note on testing latent variable models. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01715

Sofologi, M., Papantoniou, G., Avgita, T., Lyraki, A., Thomaidou, C., Zaragas, H., Ntritsos, G., Varsamis, P., Staikopoulos, K., Kougioumtzis, G., Papantoniou, A., & Moraitou, D. (2022). The Gifted Rating Scales-Preschool/Kindergarten Form (GRS-P): A Preliminary Examination of Their Psychometric Properties in Two Greek Samples. *Diagnostics (Basel, Switzerland), 12*(11), 2809. https://doi.org/10.3390/diagnostics12112809

Spanos. A. (2019). *Probability Theory and Statistical Inference: Empirical Modeling with Observational Data, Cambridge University Press*, Cambridge

Spanos, A. (2021). *Modeling vs. inference in frequentist statistics: ensuring the tustworthiness of empirical evidence*. Descargado desde: https://errorstatistics.files.wordpress.com/2021/03/modeling-vs-inference-3-2021.pdf

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modelling. *Personality and Individual Differences*, 42(5), 893-898.

Tourón, M., Navarro-Asensio, E., Tourón, J. (2023). Validez de Constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres, (GRS 2), en España. *Revista de Educación, 1*(402), 55-83. https://doi.org/10.4438/1988-592X-RE-2023-402-595

Wulf, J. N., Sajons, G. B., Pogrebna, G. et al. (2023). Common methodological mistakes. *The Leadership Quarterly, 34*(1), 101677, https://doi.org/10.1016/j.leaqua.2023.101677

Información de contacto: José A. Martínez Universidad Politécnica de Cartagena. Facultad de Ciencias de la Empresa. Departamento de Economía de la Empresa. Calle Real, 3, 3021, Cartagena. E-mail: josean.martinez@upct.es

Recibido: 14/12/2023 Aceptado: 01/06/2024

Réplica a la "Crítica del análisis de la validez de constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GRS 2); réplica a Tourón et al. (2023)"

Reply to "Criticism of the analysis of Construct Validity of the Gifted Rating Scales (GRS 2) Parent Form in Spain; a reply to Tourón et al. (2023)"

https://doi.org/10.4438/1988-592X-RE-2024-406-638

Marta Tourón
https://orcid.og/0000-0001-5430-4198
Universidad Internacional de La Rioja - UNIR

Enrique Navarro-Asencio
https://orcid.og/0000-0002-3052-146X
Universidad Complutense de Madrid

Javier Tourón
https://orcid.og/0000-0001-8217-1556
Universidad Internacional de La Rioja - UNIR

Resumen

El objetivo de este trabajo es responder a la revisión crítica realizada por Martínez, J. A. (2024) del estudio de la validez de constructo de la Escala GRS 2 para padres en España, Tourón et al. (2023). La crítica se centra principalmente en el modelo factorial propuesto, que utiliza una aproximación reflectiva del Análisis Factorial Confirmatorio (AFC) para la medida de la percepción de las altas capacidades. Mientras que en la revisión el autor propone una aproximación formativa. En esta respuesta se lleva a cabo una aclaración conceptual del término *giftedness* y también se realiza una réplica de las observaciones realizadas a la metodología de validación, poniendo la atención en las diferencias entre modelos reflectivos y formativos y las características del AFC utilizado. Además, se aportan

Recibido: 27/05/2024 Aceptado: 07/06/2024

evidencias que justifican la definición de un modelo de medida reflectivo a través de la estimación de un modelo bifactorial. Este modelo trata de explicar las respuestas a los ítems a partir de la consideración, al mismo tiempo, de un factor general y un conjunto de factores específicos. Los resultados muestran la importancia de un factor general, pero sin evidencias consistentes de una unidimensionalidad estricta. La presencia de una estructura multidimensional con cuatro factores específicos (capacidad cognitiva, capacidad creativa, habilidades sociales y control emocional) aporta una parte importante a la explicación de la varianza común del modelo.

Palabras clave: escala de detección, altas capacidades, validez de constructo, análisis factorial confirmatorio, modelos reflectivos y formativos, modelo bifactorial.

Abstract

The objective of this work is to respond to the critical review carried out by Martínez, J. A. (2024) of the study on the construct validity of the Gifted Rating Scales (GRS 2) parent form in Spain, Tourón et al. (2023). The critique mainly focuses on the proposed factorial model, which uses a reflective approach of Confirmatory Factor Analysis (CFA) for measuring the perception of high abilities, while the author in the review proposes a formative approach. This response provides a conceptual clarification of the term giftedness and also replicates the observations made on the validation methodology, focusing on the differences between reflective and formative models and the characteristics of the CFA used. Additionally, evidence is provided to justify the definition of a reflective measurement model through the estimation of a bifactorial model. This model seeks to explain the responses to the items by considering both a general factor and a set of specific factors simultaneously. The results show the importance of a general factor, but without consistent evidence of strict unidimensionality. The presence of a multidimensional structure with four specific factors (cognitive ability, creative ability, social skills, and emotional control) significantly contributes to the explanation of the common variance of the model.

Keywords: gifted rating scales, high ability, construct validity, confirmatory factor analysis, reflective and formative models, bifactor model.

Introducción

Agradecemos el minucioso estudio y análisis de Martínez, J. A. a nuestro trabajo de validación de constructo sobre la Escala de detección para Padres de las GRS 2 (*Gifted Rating Scales 2*) en España, en la que se

critican «varios de los procedimientos metodológicos y de análisis de resultados descritos por Tourón et al. (2023) son cuestionables. Tanto el modelo factorial propuesto, basado en una visión reflectiva de la medición de las altas capacidades, como los índices de ajuste empleados para validar la estructura dimensional, de carácter aproximado, limitan seriamente la interpretación de los resultados».

Vamos a replicar a las observaciones de Martínez, J. A. siguiendo la secuencia de su propia crítica y aportando nuestros argumentos respecto al trabajo realizado.

Este texto trata de dar respuesta a la crítica en diferentes aspectos relacionados con el modelo y la metodología utilizada en Tourón et al. (2023) para validar la estructura de dimensiones de la GRS 2 en su versión para familias. Recordemos que este instrumento es una escala de detección de comportamientos dotados basada en un modelo multidimensional de las altas capacidades, concretamente tres dimensiones: capacidades cognitivas, capacidades creativas y artísticas y habilidades socioemocionales (Pfeiffer y Jarosewich, 2003). Los ítems incluidos, un total de 20, reflejan comportamientos o características indicativas de alta capacidad y que pueden ser observadas por un padre o madre fuera de un entorno educativo. La validación no puede entenderse sin la consideración del uso de las puntuaciones que produce el test. Y, en este caso, la escala GRS 2 para padres, proporciona información que complementa la aportada por los docentes para identificar esos comportamientos. Antes de comenzar con la respuesta propiamente dicha, conviene proporcionar algo de contexto.

Sobre el constructo y la terminología

Hay una cuestión terminológica previa que conviene precisar. La llamada *giftedness* en inglés –que es el idioma en el que se ha escrito la literatura más relevante sobre el tema– significa dotación, de modo que el *gifted* es el dotado. El uso que se le ha dado a estos términos en castellano, en la mayor parte de las ocasiones, ha sido *superdotación* y *superdotado*. Términos no del todo correctos y no correspondientes con la literatura al respecto. Por ello, Tourón (2023) ha propuesto repetidamente atender a la dimensión sustantiva del término y utilizar las palabras *alta capacidad*. De modo que la capacidad es la dimensión sustantiva de la *giftedness*.

Muchos autores actuales (ver p. e. los planteamientos de Renzulli, Gagné o Subotnik et al. en Pfeiffer et al., 2018) entienden la *giftedness* como un constructo social, como una denominación que se refiere a un constructo multidimensional que tiene manifestaciones conductuales diversas en los planos cognitivo, afectivo, social, etc.

Tourón (2020) señalaba que «el problema conceptual que se da en este campo de estudio es mayúsculo y tiene consecuencias indeseables para la identificación y la intervención, lo que no es exclusivo de nuestro país. En cierta medida, ocurre en el ámbito anglosajón dentro de la comunidad educativa, no investigadora. Una primera problemática se deriva de la equiparación de términos como *gifted*, que quiere decir dotado, y hacerlo equivaler a *superdotado* (su correlato en ingles sería *supergifted*, que no existe), y este último ligarlo a la obtención de un valor de CI, generalmente 130, en alguno de los tests al uso. El problema central aquí es que la etiqueta se hace equivaler a un "estado del ser", o a un rasgo o cromosoma de oro (Cf. Renzulli y Reis, 2018), de manera que unas personas lo poseen y otras no. Se convierte de este modo el hecho de ser *gifted* en una variable dicotómica, cuyas categorías se establecen arbitrariamente a partir de un punto de corte (...). "El término 'dotado' [gifted en el original] significa que uno es excepcional en algo y preferimos usar la palabra "dotado" como adjetivo (por ejemplo, él o ella es un/a pianista, escritor/a, etc. dotado/a), en vez de como un sustantivo (ella es dotada). También preferimos hablar de comportamientos dotados (adjetivo) en lugar de usar los 'dotados' para representar un estado de ser (p.185)" (...).

» Todos los modelos actuales (...) enfatizan la importancia del desarrollo a lo largo de la vida de la persona, estableciendo la relevancia del impacto del entorno en dicho desarrollo. (...). Por ello, parece sensato entender que esta conceptualización nos ha de poner en primer plano la necesidad de identificar los potenciales diversos de las personas para ayudarlas a convertirlos en talentos o competencias desarrolladas, en rendimiento (Cf. Tourón, 2012). Más aún, como señala Pfeiffer (2017), el término *gifted* es un constructo social, una denominación que utilizamos para referirnos a un grupo heterogéneo de personas que se caracterizan por tener una alta capacidad, un alto rendimiento o potencial para rendir» (p.17-18).

Es necesario precisar que las escalas GRS 2 (en nuestro caso la de padres, aunque hay otras dos de profesores) proceden de una concepción previa de la *giftedness,* como realidad multidimensional que

Recibido: 27/05/2024 Aceptado: 07/06/2024

tiene manifestaciones externas en la conducta o comportamiento de las personas. Este conjunto de características observables, típicas de las personas con altas capacidades, se traducen en ítems que los padres (en este caso) han de valorar. De este modo, debería esperarse que los ítems diseñados para medir una determinada dimensión se relacionen entre sí de manera más intensa que si se han elaborado para medir dimensiones distintas. De esta forma, las dimensiones del constructo se consideran variables o dimensiones latentes, causantes de las respuestas a los ítems, que se consideran efectos.

Así, la escala es la dimensión operativa, medible, aunque sea de modo imperfecto, que permite estimar los rasgos latentes no medibles. Esto es un enfoque claramente reflectivo. La concepción de lo medido es anterior y base del instrumento de medida.

Finalmente, es preciso señalar que esta escala, como muchas otras en su género, no pretende medir las *altas capacidades*, como ocurre con un test de capacidad, por ejemplo, sino que pretende recabar información de otras fuentes (los padres) que valoran el grado en el que determinadas conductas o características están presentes –y en qué grado– en la persona valorada.

La estructura de la escala y su construcción es posterior a la conceptualización del constructo que pretende valorar, pero no es un instrumento que pretenda definir qué es la *giftedness*. (ver p. e. Pfeiffer y Jarosewich, 2003; Tourón et al., 2024).

En el trabajo de Tourón et al. (2023) los autores se centran en aportar evidencias psicométricas de la validez de constructo de la traducción española de la escala, pero no han participado en la definición teórica y operativa del constructo. También debe considerarse que la respuesta a los ítems que incluye la escala es producto de la percepción que tienen las familias sobre determinados comportamientos de sus hijos. Y que la respuesta a cada uno de esos ítems tiene un formato ordinal, utilizando una escala Likert de 6 puntos para identificar la frecuencia de cada uno de esos comportamientos.

Sobre la validación

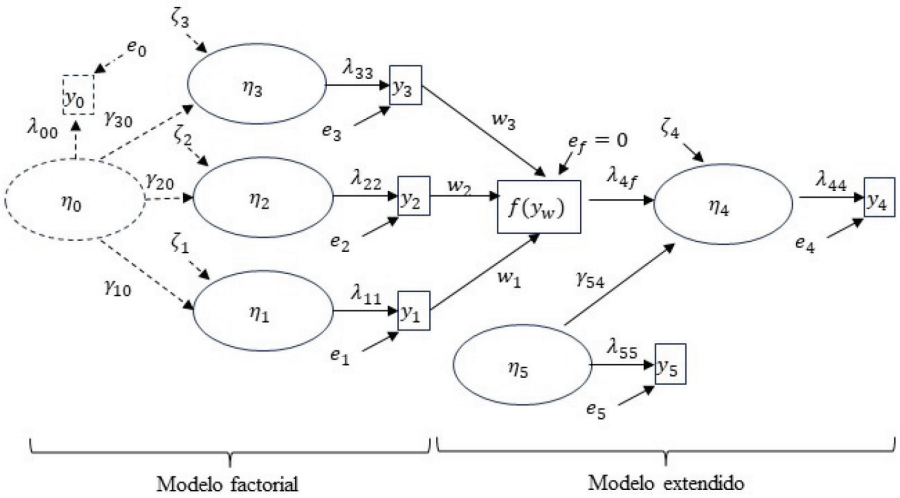
La validación de instrumentos tiene el propósito de recoger evidencias que apoyen la interpretación y uso de las puntuaciones de los test, como se señala en el manual de AERA, APA y NCME (2018). La validez se

refiere, como se sabe, al grado en el que la evidencia y la teoría apoyan las interpretaciones de las puntuaciones de los test asociadas a los usos previstos (p. 11). Las evidencias de validez pueden obtenerse de cinco fuentes: el contenido del test, el proceso de respuesta, la estructura interna, las relaciones con otras variables y las consecuencias de la aplicación del test. En el trabajo de Tourón et al. (2023) se aportan evidencias de la validez de la estructura interna de los ítems que forman la escala y su organización en dimensiones, pero tampoco es su propósito estimar las puntuaciones en los factores o baremar el instrumento.

La crítica metodológica se centra principalmente en la definición del modelo de medida utilizado para recoger evidencias de la relación entre los ítems y los constructos latentes. En el modelo confirmatorio utilizado en nuestro trabajo las respuestas a los ítems de la escala son considerados variables dependientes, es decir, son los efectos producidos por los factores latentes, que en nuestro modelo son las variables exógenas, las causas. Sin embargo, en la revisión se propone invertir el papel de ítems y factores, considerando los constructos formativos y no reflectivos.

En la revisión crítica, el autor argumenta que la aproximación más ajustada para definir la forma de calcular las puntuaciones factoriales es un modelo formativo, donde las puntuaciones en los diferentes ítems determinan el factor. De forma opuesta a la aproximación reflectiva utilizada en el trabajo, como se muestra en el modelo extendido de la siguiente figura:

FIGURA I. Modelo factorial y modelo extendido



Fuente: Martínez, J. A. (2024)

La parte izquierda de la figura, el modelo factorial, representa nuestra estimación y el modelo de la derecha la propuesta de la crítica que se nos hace.

La utilización de una definición u otra es un asunto controvertido en la literatura psicométrica (Murray y Booth, 2018), optar por uno de los modelos implica asumir supuestos que van desde la definición teórica y operativa del constructo, hasta su validación empírica. Emplear una aproximación formativa conlleva tratar los factores latentes como variables endógenas (variable dependiente), es decir, como un efecto producto de la combinación de diferentes indicadores. Y, por tanto, no considerarlos variables independientes, un factor exógeno que determina cómo los sujetos responden a los ítems.

Murray y Booth (2018) mencionan 5 aspectos para diferenciar entre indicadores como causas o efectos. El primero, es la dirección de causalidad; en los modelos formativos, los ítems son la causa del factor latente y en los reflectivos son los efectos, las variables a explicar. El segundo, son los cambios que los indicadores pueden producir en el constructo; en el modelo formativo cualquier cambio en el indicador producirá un efecto en el valor del constructo, por el contrario, en los reflectivos los cambios en un indicador no deben altear el nivel en el constructo. El tercer argumento se refiere a si el constructo es la causa común del conjunto de indicadores, cuestión que se produce en el caso del modelo reflectivo, donde los indicadores necesitan estar correlacionados. El cuarto, es si los indicadores tienen las mismas causas y efectos, un requisito que es necesario en los modelos reflectivos, pero no en los formativos. Finalmente, también mencionan la intercambiabilidad de los indicadores que, en principio, debe cumplirse en los modelos reflectivos.

Siguiendo a Zumbo (2006), la distinción reflectiva y formativa diferencia entre medidas e índices para tratar a los indicadores o variables observadas. En el primer caso, un cambio en la variable latente se refleja en cambios en los indicadores (las respuestas a los ítems que se consideran variables dependientes o endógenas). Y, en el segundo, el índice es la variable latente que provoca cambios en el factor, son, por tanto, las causas (variables independientes o exógenas). Este autor señala que realizar un Análisis de Componentes Principales (ACP) no es evidencia suficiente para validar un constructo desde una aproximación formativa y tampoco los procedimientos de Análisis Factorial Confirmatorio (AFC) determinan que se cuente con un constructo reflectivo. En el trabajo de Bollen y

Lennox (1991) proponen que esta diferenciación debe explorarse durante la validación de contenido, preguntando a los expertos por el carácter formativo o reflectivo de la medida o utilizando técnicas de pensamiento en voz alta para la recogida de evidencias durante una aplicación piloto.

¿Por qué un modelo reflectivo?

El modelo utilizado en Tourón et al. (2023) busca aportar evidencias psicométricas de la estructura de dimensiones de la escala y también de la calidad de los indicadores utilizados para su medida. En este contexto, el trabajo estudia distintos modelos para contrastar empíricamente diferente número de dimensiones y propone un modelo final de cuatro dimensiones, aunque también se prueba y muestra un buen ajuste un modelo con dos factores de segundo orden. Desde esta perspectiva, se asume la existencia de constructos no observables que determinan la forma de responder a los ítems que incluye un test. Y también que en esa respuesta hay parte explicada por el constructo latente y otra parte que es error de medida.

Debe tenerse en cuenta que el trabajo de Tourón et al. (2023) valida la estructura de dimensiones de la traducción del instrumento GRS 2, pero no han participado en el análisis inicial de la validez del contenido, que tiene el propósito de validar el conjunto de indicadores y asegurar que son una muestra de conductas de la red nomológica que define el constructo. Este tipo de validez suele realizarse a partir de un juicio de expertos. Aunque se asume que los ítems de la escala GRS 2 cumplen con la validez de contenido. y qué el propósito de los indicadores es representar las conductas del constructo de medida, pero que pueden hacer referencia a diferentes niveles del mismo factor.

La descripción hecha en la revisión crítica debe completarse con las siguientes puntualizaciones.

En primer lugar, Martínez, J. A. menciona que "en línea con la teoría clásica de los tests: la variación en la variable latente η_1 se manifiesta en una variación en el indicador observable, escalada con el parámetro λ_11, a lo que hay que añadir un error aleatorio e_1." (p. 5) De la siguiente forma:

$$y_1 = \lambda_{11}\eta_1 + e_1 \qquad (1)$$

Recibido: 27/05/2024 Aceptado: 07/06/2024

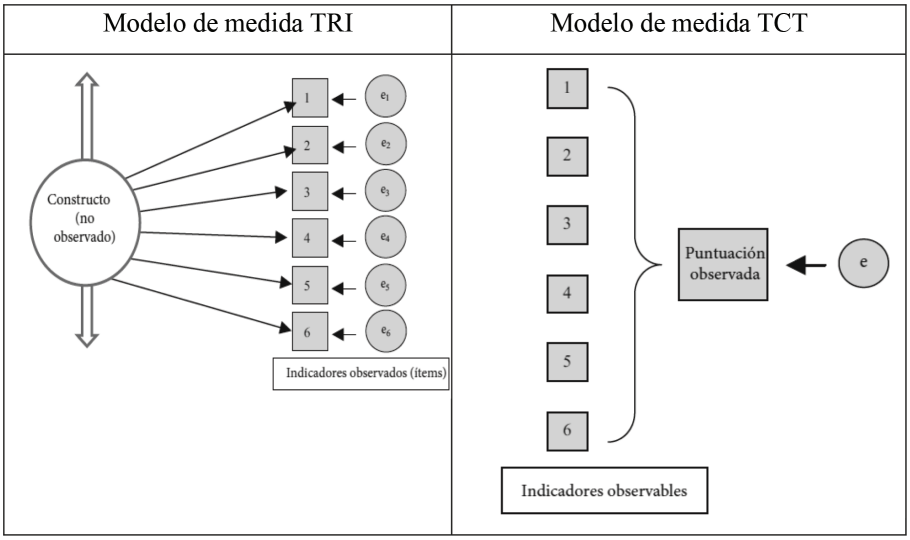
Sin embargo, esta aproximación para definir las respuestas a los ítems se ajusta más a los supuestos de la TRI que de la TCT. En la TCT o Teoría de la puntuación verdadera, el modelo pone la atención en la definición de la puntuación total en el test como resultado de la suma de la puntuación verdadera y un error de medida:

$$X=V+E \qquad (2)$$

La dispersión de esos errores es un indicador de la precisión de las puntuaciones observadas y puede estimarse mediante la fiabilidad. Las inferencias se hacen utilizando la puntuación total en el test (escala), que se calcula a partir de la información observada. En cambio, en la TRI el modelo define de forma explícita el comportamiento del constructo a medir, del rasgo latente. Son las respuestas a los ítems las que están determinadas, en términos de probabilidad, por el nivel que tienen un sujeto en el constructo.

El modelo de la TCT no incluye información sobre la relación entre el constructo latente y la forma de responder a los ítems. Solo el concepto de puntuación verdadera hace esa referencia, es parte no observada. La diferencia entre los dos modelos de medida puede verse en el siguiente gráfico:

FIGURA II. Modelos de medida en la TCT y la TRI



Fuente: Adaptado de Wu, Tam y Jen (2016).

En el modelo TRI las flechas señalan el efecto del constructo en la probabilidad de responder de una forma u otra al ítem. Es, por tanto, el constructo el que determina todo el patrón de respuestas (modelo reflectivo). Los errores (e) también se representan con círculos, son variables que no se observan directamente, y señalan la influencia en la respuesta de otros factores desconocidos, distintos al constructo que queremos medir. El error es un término asociado a cada ítem y determina la capacidad que tiene el constructo para explicar la respuesta al mismo.

En la TCT, tanto las respuestas a los ítems como la puntuación total del test son datos observados. El total del test se calcula como una agregación de las respuestas a los ítems (suma de aciertos, promedios, etc.) (modelo formativo) y el error está asociado a esta puntuación total y no a cada uno de los ítems. La diferencia entre estos dos modelos refleja la aproximación reflectiva y formativa utilizada para definir el constructo y también las diferencias entre un Análisis Factorial Confirmatorio (AFC) y un Análisis de Componentes Principales (ACP).

En segundo lugar, Martínez, J. A. presenta una aproximación lineal del análisis factorial, y la perspectiva utilizada en el trabajo de Tourón et al. (2023), debido al carácter ordinal de los ítems y la falta de normalidad multivariada, es no lineal. Así, el modelo para definir la respuesta a los ítems estima la puntuación en el constructo necesaria para situarse entre las diferentes opciones de respuesta. Se empleó la matriz de correlaciones policóricas como elemento informativo en el análisis factorial. Al utilizar este tipo de correlación se asume la existencia de una variable continua subyacente (Jöreskog, 1994) y las respuestas politómicas observadas se consideran manifestaciones de los encuestados que superan un determinado número de puntos de corte o umbrales dentro de ese continuo. En este sentido, el modelo estima esos umbrales (*thresholds* "τ") y define las respuestas observadas en las diferentes categorías ordinales a través de variables continuas latentes. Concretamente, para un ítem i con un número de categorías c=0,1,2…, C la variable latente y* se define de forma que:

$$yi = c \text{ if } \tau c < yi^* < \tau c+1 \qquad (3)$$

donde τc, τc+1 son los umbrales que determinan los puntos de corte en la variable continua latente subyacente, que suelen estar espaciados en intervalos de diferente amplitud. Considerando este supuesto, la correlación de interés para el modelo está entre estas variables continuas

Recibido: 27/05/2024 Aceptado: 07/06/2024

(correlación policórica). El procedimiento de análisis suele realizarse en tres pasos (Jöreskog,1990; Muthén, 1984). En las primeras dos etapas se estiman los umbrales y las correlaciones policóricas y, en la tercera etapa, estos valores se ajustan en un modelo hipotetizado utilizando algún método de estimación, principalmente DWLS. Los parámetros del modelo se obtienen minimizando la función que compara la información estimada con los datos del modelo.

En tercer lugar, en la revisión Martínez, J. A. menciona una variable latente de orden superior, a modo de factor general, «conceptualización habitual en esta área de conocimiento es la primera, tal y como se muestra, por ejemplo, en Pfeiffer et al. (2008), donde se plantea una concepción multidimensional de las altas capacidades con un factor subyacente "g" o factor de capacidad general. Hay que reconocer que Tourón et al. (2023) son muy cautos y no afirman claramente este hecho, pero subrepticiamente parecen hacerlo cuando calculan una "media total de la escala", como posteriormente explicaremos» (p. 6). Aquí es conveniente mencionar que en la validación no se prueba en ningún momento un factor general, sino que se presenta una estructura con dos factores de segundo orden. El primero determina las puntuaciones en el factor cognitivo y el factor creativo y, el segundo, para el factor social y el factor de control emocional. Además, la media del conjunto de ítems calculada en la sección de estadísticos descriptivos no pretende mostrar la existencia de ese factor general, es solo un resumen de los datos. También conviene recordar que el objetivo del trabajo consistió en aportar evidencias de la validez de constructo, pero no estimar las puntuaciones en las dimensiones.

En cuarto lugar, relacionado con lo anterior, Martínez, J. A. menciona que «proponer un modelo multidimensional (…) que refleja un factor g subyacente, implica desde el punto de vista de la medida, que el factor g puede medirse con el "mejor" indicador de la "mejor" dimensión, es decir, con un único ítem(…) Esto es compatible, como no podría ser de otro modo, con la visión reflectiva sobre la medición, donde los ítems de una variable latente se pueden considerar intercambiables, lo que indica que quitando un ítem no se altera el significado de la variable latente» (p. 7).

En este argumento se hace referencia a la intercambiabilidad de los indicadores utilizados para definir el constructo y también para la definición de los factores de segundo orden. No queda claro si Martínez, J. A. alude aquí al supuesto de paralelismo de las puntuaciones de

la TCT (Lord y Novick, 1968) o el paralelismo conceptual de los indicadores que definen el constructo, que asume que intercambiar uno por otro no altera el significado del constructo (Borsboom et al., 2004). En el primer tipo de paralelismo, para estimar la fiabilidad de un test con datos empíricos, los ítems se consideran pequeñas partes paralelas que reflejan la puntuación verdadera y se utilizan para la evaluación de la consistencia interna. Lograr medidas completamente paralelas implica que las medias, las desviaciones típicas y errores de medida son equivalentes, algo muy complejo de lograr en la práctica. Si no se logra la misma dispersión nos situamos ante medidas tau-equivalentes. Otro caso de paralelismo, las medidas esencialmente tau-equivalentes, permite también que las medias varíen entre las diferentes partes añadiendo una constante. Finalmente, las medidas congenéricas son las menos restrictivas y permiten también que esas medias difieran añadiendo o multiplicando por una constante. En nuestra validación, no puede asumirse una estructura completamente paralela porque los ítems de un mismo constructo son una muestra de los comportamientos, pero pueden reflejar diferentes intensidades de esa medida. En cualquier caso, los modelos de AFC permiten contar con medidas que no cumplen con el paralelismo estricto, es decir, con pesos factoriales distintos. Además, la base de este análisis es la consistencia interna de los indicadores que definen el constructo, donde se espera que haya correlación. Sin embargo, la aproximación formativa en un ACP no necesita que los indicadores que determinan el factor estén correlacionados, además se asume que no son intercambiables. El AFC permiten contar con indicadores que tienen una correlación heterogénea con los factores latentes. Aquí, la fiabilidad compuesta (ω) se calcula a partir de las cargas factoriales para producir estimaciones más precisas que los procedimientos de consistencia interna, como el α de Cronbach.

En los factores de segundo orden sí puede ser más ajustado definir un modelo formativo si los diferentes factores asociados a la alta capacidad se consideran partes de un factor general. Considerando los resultados de la validación, los dos factores de segundo orden (Capacidades Cognitivo-Creativas y Habilidades Socio-Emocionales) definidos no pueden considerarse intercambiables. Como muestran los resultados de correlación (0,53), comparten el 25% de variabilidad. No obstante, como ya hemos mencionado, el factor general no se probó en nuestra validación.

Recibido: 27/05/2024 Aceptado: 07/06/2024

En quinto lugar, Martínez, J. A. realiza una analogía entre capacidades intelectuales y físicas. Aquí es conveniente apuntar que la escala GRS es un instrumento de percepción, no mide directamente la capacidad intelectual, sino que se infiere a partir de la observación de determinados comportamientos. Además, los indicadores para medir capacidades físicas son observables y sin error, a no ser que el instrumento utilizado para recoger la medida no funcione correctamente. En cambio, la percepción estará determinada en parte por el constructo, pero también puede estar influida por algún factor no medido (error).

Es posible que para la medida de la capacidad física sea necesaria una única medida de cada indicador y una combinación de todos ellos para estimar la dimensión. Sin embargo, para poder medir con precisión dimensiones latentes a partir de indicadores basados en la respuesta a ítems, es necesaria más de una medida de la misma conducta. Por ejemplo, para medir la capacidad aritmética en competencia matemática de Educación Primaria, puede incluir un único ejercicio para sumar dos cantidades o incluir varios con el mismo propósito en un mismo examen. Más ítems aumentarán la fiabilidad de la medida.

En sexto lugar, Martínez, J. A. también propone reducir el número de indicadores para probar esta cuestión, sin embargo, como ya hemos mencionado, los indicadores pueden reflejar diferentes niveles del constructo al no considerarse medidas paralelas y también dependerá de la complejidad conceptual del constructo medido. Insistimos en que los indicadores pueden considerarse intercambiables, en el modo que son conductas determinadas por el mismo constructo, pero para poder medir con mayor precisión son necesarios varios indicadores. En caso contrario, las puntuaciones tendrían mucho error de medida.

Martínez, J. A. también menciona que «Tourón et al. (2023) parecen considerar como algo positivo cuando aluden a las correlaciones altas entre cada ítem y el resto señalando la *"homogeneidad del conjunto de datos"*. Es más, hay ejemplos en la propia literatura sobre altas capacidades y escalas GRS donde se vislumbra un efecto halo que afecta obviamente a la validez (ej. Jabůrek et al., 2020).» (p. 12). No obstante, asumiendo la posible causa común de las respuestas a los indicadores y la consistencia interna de los mismos, esta homogeneidad es una característica de los modelos reflectivos. La respuesta a los indicadores está provocada, en parte, por el factor latente, pero el modelo asume error de medida. Ese error muestra que hay otros factores que pueden

determinar esa respuesta. Martínez, J. A. también menciona que «lo que probablemente está sucediendo es que esa batería de indicadores está midiendo variables latentes diferentes, es decir, no son la manifestación de una única variable latente, sino de varias.» (p. 12). En este sentido, considerando los valores de r^2 de los ítems y los promedios de varianza explicada en cada factor (AVE), la parte explicada por las dimensiones latentes es mayor que el posible efecto de otros factores no tenidos en cuenta. No obstante, aportar evidencias de variables que puedan determinar el sesgo podría ser otra evidencia más de la validez de constructo.

Finalmente, en séptimo lugar, Martínez, J. A. realiza una crítica al procedimiento seguido en la metodología, incluyendo la utilización de un Análisis Factorial Exploratorio, como etapa previa al AFC y los índices de ajuste utilizados. Se citan los trabajos de Hayduk (2014a y 2014b) como justificación de la crítica, mencionado que «el análisis factorial exploratorio es incapaz de detectar la estructura real de los datos, es decir, de identificar el modelo que ha generado esos datos empíricos» (p. 11), argumentando que χ^2 puede utilizarse incluso con tamaños de muestra grandes y que si no pasa este test no pueden interpretarse el resto de los parámetros. Es conveniente mencionar aquí que el AFE, realizado como primera etapa, no tiene un propósito confirmatorio, se emplea para extraer información inicial sobre el número de dimensiones y recoger evidencias de la consistencia de la estructura de dimensiones con el modelo confirmatorio. En la validación de la escala GRS 2, se prueban diferentes modelos confirmatorios, de tres y cuatro dimensiones y se toman decisiones considerando los índices de modificación, por ejemplo, el cambio del ítem 17 del factor cognitivo al factor creativo.

La afirmación sobre el índice de ajuste χ^2 es demasiado contundente a nuestro parecer. Los procedimientos de validación de constructo en el ámbito de la medición educativa y la psicometría apuntan la necesidad de interpretar con precaución este índice de ajuste en el contexto de los AFC. El estadístico χ^2 pone a prueba la hipótesis nula de que la matriz de covarianzas observadas en la población es equivalente a la producida por el modelo, sin embargo, en Ciencias Sociales, cualquier modelo se considera una aproximación a la realidad y, por tanto, esa hipótesis nula con un ajuste exacto es inviable (Bentler y Bonett, 1980; Schermelleh-Engel et al., 2003). Aunque Hayduk (2014b) indica que los valores de χ^2 no están afectados por el tamaño muestral cuando el modelo está

correctamente especificado, pero considerando esa particularidad de los modelos utilizados en Ciencias Sociales no sería posible. El mismo autor apunta que aumentar la muestra incrementa el poder de χ^2 para detectar problemas en la especificación del modelo. La cuestión aquí es saber si, cuando el valor de χ^2 resulta significativo, es realmente un indicador suficiente para descartar el modelo y poder cometer un error Tipo I o si ese modelo es la mejor aproximación a la realidad y el posible sesgo asumible. Wang y Wang (2020) argumentan que la probabilidad de rechazar un modelo incrementa sustancialmente cuando el tamaño de las muestras aumenta, incluso cuando las diferencias entre las matrices de varianzas-covarianzas observadas y estimadas son pequeñas. Señalan que el valor de χ^2 aumenta al no cumplir con el supuesto de normalidad multivariada y las distribuciones de respuesta a los ítems son asimétricas o están afectadas por algún tipo de curtosis. Y también que el índice χ^2, cuando el número de variables en el modelo se incremente. Razones suficientes para no utilizarlo de forma exclusiva.

En nuestro caso, la estimación utiliza una ponderación de la matriz de varianzas y covarianzas para los valores de χ^2 y los errores típicos. Recordemos que, al no cumplir con el supuesto de normalidad multivariada, se utilizó WLSMV y la matriz de correlaciones policórica. Este método de estimación es una versión de DWLS (Muthén, 1984), pero aplicando la corrección robusta de media y varianza a los mínimos cuadrados ponderados y un cambio en la escala (son los denominados WLSMV o ULSMV). Y, aunque en el trabajo se presentan los valores de χ^2, con este tipo de datos no conviene utilizarlo (Finney y DiStefano, 2013). En el trabajo de Shi et al. (2018) se estudió cómo afecta el tipo de estimación, el tamaño de la muestra o la complejidad del modelo a diferentes índices de χ^2 y muestran un mejor funcionamiento general de los índices robustos.

Bentler y Bonett (1980) propusieron el uso de índices de ajuste incremental, como CFI y TLI, para calcular la cantidad de información que se gana cuando se comparan modelos. Estos índices evalúan el grado en el que un modelo estimado es mejor que un modelo nulo (modelo en el que todas las variables observadas no están correlacionadas) en su capacidad para reproducir la matriz de varianzas-covarianzas observadas. Y, sobre todo, las medidas de ajuste absoluto, donde se comprueba si el modelo definido se corresponde con los datos empíricos. Los índices de ajuste absoluto, como RMSEA o SRMR, no utilizan un modelo nulo

de referencia, pero realizan una comparación implícita con un modelo saturado que reproduce de forma exacta la matriz de varianzas-covarianzas de las variables observadas (Hu y Bentler, 1999). El índice RMSEA señala la falta de ajuste entre el modelo especificado en la población y el SRMR estima la raíz cuadrada del promedio de los residuos.

Considerando lo anterior, todas las aportaciones que permitan comprobar de forma empírica la validez del constructo de la escala GRS 2 refuerzan la calidad psicométrica del instrumento. Con el propósito de afianzar las evidencias a favor del modelo propuesto originalmente en el trabajo de Tourón et al. (2023), se ha llevado a cabo un modelo bifactorial (Holzinger y Swineford, 1937; Chen et al., 2006) para comprobar la existencia de un factor general que pueda ser una causa común de las respuestas a los ítems, junto con los cuatro factores definidos.

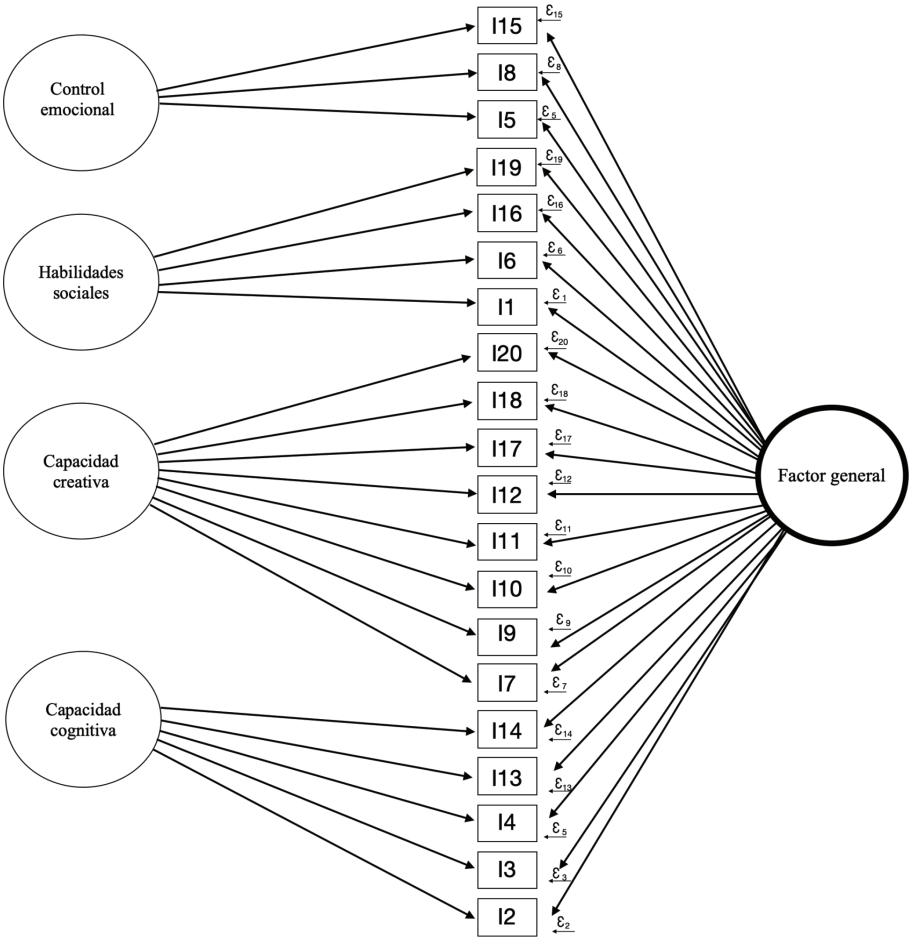
Modelo bifactorial

Esta propuesta define que la varianza compartida entre las respuestas a los ítems en dos partes, por un lado, la explicada por un factor general y aquella que está determinada por un grupo de factores específicos, que pueden ser del mismo dominio. Por tanto, plantea la hipótesis de un factor general para explicar la varianza común y, al mismo tiempo, también la existencia de múltiples factores que tienen un impacto independiente en la explicación de esa varianza (ver Figura III). Dado que los factores específicos se interpretan como la varianza contabilizada por encima y más allá del factor general, se asume que las relaciones entre los factores generales y específicos son ortogonales (no correlacionadas).

Chen y Zhang (2018) señalan que la capacidad de estudiar los factores específicos independientemente del factor general es importante para comprender mejor las afirmaciones teóricas. Por ejemplo, si un factor específico propuesto no explicara una cantidad sustancial de varianza más allá del factor general, se observarían cargas factoriales pequeñas y no significativas en el factor específico, así como una varianza no significativa del factor específico en el modelo bifactorial. Cuestión que evidenciaría que el factor específico no proporciona una explicación a la varianza más allá del factor general.

Para llevar a cabo una evaluación del modelo, después de obtener el ajuste del modelo tanto para modelos unidimensionales como

FIGURA III. Modelo bifactorial



Fuente: Elaboración propia.

bifactoriales, se pueden comparar directamente mediante la diferencia entre los índices CFI (ΔCFI), ya que el modelo unidimensional está jerárquicamente anidado dentro del modelo bifactorial (Reise, 2012). Este índice se calcula como:

$$\Delta CFI = CFI\_M1 - CFI\_M0 \quad (4)$$

Donde CFI_M1 es igual al valor CFI obtenido para el modelo 1, y CFI_M0 es igual al valor CFI obtenido para el modelo 0. Este índice es más estable en condiciones diferentes, como el tamaño de la muestra, la cantidad de error, el número de factores y el número de ítems y se recomiendan valores iguales o inferiores a .01 para confirmar la equivalencia (Meade et al., 2008).

Para la fiabilidad se pueden emplear cuatro versiones del coeficiente omega (o fiabilidad compuesta): omega total para el factor general (ωt), omega para cada subdimensión (ωs) y los coeficientes omega jerárquicos, que también pueden calcularse para el factor general (ωht) y para cada subdimensión (ωhs). El coeficiente omega McDonald (1999) es un tipo de fiabilidad basada en los resultados del análisis factorial, adecuado cuando se cuenta con medidas congenéricas (distintos pesos factoriales), que estima la proporción de la varianza observada atribuible a los factores del modelo:

$$\omega t = \frac{(\sum \lambda_G)^2 + (\sum \lambda_{Cog})^2 + (\sum \lambda_{Crea})^2 + (\sum \lambda_{Soc})^2 + (\sum \lambda_{Emo})^2}{(\sum \lambda_G)^2 + (\sum \lambda_{Cog})^2 + (\sum \lambda_{Crea})^2 + (\sum \lambda_{Soc})^2 + (\sum \lambda_{Emo})^2 + (\sum_{i=1}^{n} e_i)} \quad (5)$$

En el denominador están todas las fuentes de varianza del modelo, la varianza común producida por todos los factores del modelo, general y subdimensiones, más la varianza específica de los errores. El numerador incluye solo esas las fuentes de varianza común. También puede calcularse para cada subdimensión:

$$\omega s = \frac{(\sum \lambda_G)^2 + (\sum \lambda_{Subdimension})^2}{(\sum \lambda_G)^2 + (\sum \lambda_{Subdimension})^2 + (\sum_{i=1}^{n} e_i)} \quad (6)$$

En este caso, los pesos factoriales y los errores son de los ítems correspondientes a la subdimensión. Y para determinar la proporción de la varianza total de las puntuaciones, debida únicamente al factor general, se utiliza el coeficiente omega jerárquico, que se calcula dividendo la suma de las cargas factoriales del factor general al cuadrado por la varianza total, considerando varianza común y error:

$$\omega h = \frac{(\sum \lambda_G)^2}{(\sum \lambda_G)^2 + (\sum \lambda_{Cog})^2 + (\sum \lambda_{Crea})^2 + (\sum \lambda_{Soc})^2 + (\sum \lambda_{Emo})^2 + (\sum_{i=1}^{n} e_i)} \quad (7)$$

La varianza debida a las subdimensiones se considera aquí parte del error de medida, un coeficiente de 0,8 más señala que las puntuaciones pueden considerarse esencialmente unidimensionales, considerando el factor general como la principal fuente de varianza. Además, podemos saber cuánto aporta cada una de las dimensiones a la varianza común, una vez controlado el factor general (Reise, 2012).

$$\omega hs = \frac{(\sum \lambda_{Subdimension})^2}{(\sum \lambda_G)^2 + (\sum \lambda_{Subdimension})^2 + (\sum_{i=1}^{n} e_i)} \quad (8)$$

Como en el caso de la varianza total, solo se emplean los pesos factoriales y errores de los ítems que componen la subdimensión. Como referencia pueden emplearse los puntos de corte mencionados por Smits et al. (2014), donde los valores iguales o superiores a 0,3 pueden considerarse importantes, los inferiores hasta 0,2 moderados y por debajo de 0,2 bajos.

Desde esta perspectiva, el índice de varianza común explicada (ECV) se utiliza para contrastar la unidimensionalidad de las escalas a partir de las cargas factoriales del factor general y de las subdimensiones (Reise et al., 2013), de la siguiente forma:

$$ECV = \frac{(\sum \lambda_G)^2}{(\sum \lambda_G)^2 + (\sum \lambda_{Cog})^2 + (\sum \lambda_{Crea})^2 + (\sum \lambda_{Soc})^2 + (\sum \lambda_{Emo})^2} \quad (9)$$

Es, por tanto, la proporción de varianza explicada por el factor general dividida por la varianza explicada por ese factor y las subdimensiones. Valores altos señalan una gran importancia del factor general, pero establecer puntos de corte no es sencillo, pero se sugieren valores de 0,7 o superiores para considerar la unidimensionalidad (Rodríguez et al., 2016). Este indicador también se puede calcular en el nivel del ítem para identificar aquéllos donde la influencia del factor general es muy potente, autores como Stucky y Edelen (2014) proponen valores superiores a 0,8 o 0,85.

Para interpretar el índice de unidimensionalidad ECV, se recomienda calcular también el porcentaje de correlaciones no contaminadas (PUC) (Rodríguez et al., 2016). Este índice, junto al ECV, informan sobre el posible sesgo al forzar datos multidimensionales en un modelo unidimensional calcula cuántas correlaciones modera los efectos que puede tener el factor general. PUC puede definirse como el número de correlaciones no contaminadas dividido por el número de correlaciones únicas:

$$\mathrm{PUC} = \frac{\frac{I_G * (I_G - 1)}{2} - \left[\frac{I_{s1} * (I_{s1} - 1)}{2} + \frac{I_{s2} * (I_{s2} - 1)}{2} + \cdots + \frac{I_{sn} * (I_{sn} - 1)}{2}\right]}{\frac{I_G * (I_G - 1)}{2}} \qquad (10)$$

I_G es el número de ítems que cargan en el factor general, I_s es el número de ítems que cargan en cada factor específico. Cuando los valores de PUC están por encima de 0,8 los valores de ECV no son muy relevantes porque señala que las cargas factoriales del modelo unidimensional se aproximarán a las obtenidas en el factor general del modelo bifactorial. En cambio, si los valores de PUC están por debajo de 0,8, deben obtenerse valores de ECV superiores a 0,6 para considerar la unidimensionalidad. Y si los valores son muy altos (>,90), pueden obtenerse estimaciones unidimensionales insesgadas incluso cuando se obtiene un valor de ECV bajo (Reise, 2012).

Otros indicadores que puede ayudar a evidenciar la calidad del modelo es el índice H de replicabilidad del constructo. Este índice permite valorar si el conjunto de ítems que representa cada variable latente es adecuado y se calcula de la siguiente forma:

$$\mathrm{H} = \frac{1}{\left[1 + \frac{1}{\sum_{i=1}^{n} \frac{\lambda^2}{1 - \lambda^2}}\right]} \qquad (11)$$

Es, por tanto, un sumatorio de la tasa de la varianza explicada por la variable latente, dividida por el error de medida, es decir, la proporción de la variabilidad del constructo que se explica por sus propios indicadores. Los valores de 0,7 o más señalan que la variable latente

Recibido: 27/05/2024 Aceptado: 07/06/2024

está bien definida por sus indicadores y tendrá estabilidad en diferentes estudios (Rodríguez et al., 2016).

Para finalizar, se analizan las cargas factoriales del modelo bifactorial y se comparan los resultados del factor general con los obtenidos en el modelo puramente unidimensional (M3, ver Tabla I). A partir de las diferencias entre los valores de las cargas de cada ítem de los dos modelos se calcula el sesgo relativo de los parámetros (SRP):

$$\text{SRP} = \frac{\lambda_G - \lambda_{UNIDIM}}{\lambda_G} * 100 \qquad (12)$$

El promedio de este índice informa sobre el sesgo que puede cometerse al ajustar un modelo unidimensional cuando no se cumple ese supuesto. Los valores superiores al 15% señalarían ese posible sesgo (Rodríguez et al., 2016). Además, como señalan Ferrando y Lorenzo-Seva (2018), si en el modelo bifactorial el factor general acumula cargas altas y en las dimensiones, en cambio, el promedio de sus cargas no supera el 0,3, puede ser otro indicador de la estructura unidimensional de la medida.

Resultados

En primer lugar, se incluyen los valores de ajuste del modelo bifactorial comparados con algunos de los probados en el trabajo de Tourón et al. (2023). Concretamente, se presentan los resultados del modelo original (M1), el modelo unidimensional (M3), el modelo propuesto (M6) y su versión con dos factores de segundo orden (M8). Por último, el nuevo modelo bifactorial (M9).

El modelo bifactorial, considerando las diferencias de los índices de ajuste incremental y absoluto con el modelo propuesto (M6), pueden considerarse equivalentes en su capacidad explicativa. También muestra un mejor ajuste que el modelo puramente unidimensional (M3).

En segundo lugar, los valores de omega absoluto y jerárquico para el factor general y las subdimensiones se presentan en la siguiente tabla:

Los valores de omega (ω) señalan la fiabilidad de cada factor y el coeficiente jerárquico (ωh) muestra la parte atribuible a cada uno en la explicación de la varianza total. Como se observa, el valor de ωh del factor general está por debajo de 0,7 pero evidencia una aportación

TABLA I. Índices de ajuste de los modelos y diferencia entre el modelo bifactorial el propuesto en Tourón et. al. (2023)

Índices	**M1**	**M3**	**M6**	**M8**	**M9**	**Diferencia (Δ) M9-M8**
AFC	3	1				
χ^2	2048	6833	1596	1601	1760	
gl	167	170	164	165	150	
p	<,001	<,001	<,001	<,001	<,001	
χ^2/gl	12,263	40,194	9,732	9,703	11,733	
SRMR	0,086	0,154	0,074	0,074	0,081	0,007
RMSEA	0,101	0,189	0,089	0,089	0,098	0,009
CFI	0,968	0,867	0,976	0,976	0,973	-0,003
TLI	0,964	0,851	0,972	0,972	0,966	-0,006
GFI	0,978	0,91	0,983	0,983	0,981	-0,002

Fuente: Elaboración propia

TABLA II. Omega y Omega jerárquico de los factores (general y específicos)

	F. General	**Cognitivo**	**Creativo**	**Social**	**Emocional**
ω	0,941	0,880	0,921	0,826	0,789
ωh	0,669	0,427	0,566	0,495	0,509

Fuente: Elaboración propia.

relevante. También los factores aportan la explicación de la varianza total, con valores de *ω*h alrededor del 50%.

Comprando los resultados *ω* y *ω*h del factor general se observa si el 66,9% de la variabilidad está determinada por ese factor, el otro 27,2% está provocado por las diferencias de los factores específicos. El resto, un 5,9% se atribuye, por tanto, al error de medida.

En tercer lugar, el índice de unidimensionalidad ECV es de 0,71, es decir, un 71% de la varianza común está explicada por el factor general y el otro 29% por las subdimensiones. Recordemos que valores de 0,8 señalan la unidimensionalidad y cercanos a 0,7 podrían ser indicadores

Recibido: 27/05/2024 Aceptado: 07/06/2024

también de esa posibilidad. Sin embargo, el índice ECV para los ítems solo detecta uno con un valor por encima de 0,8, es el ítem 17, como muestra la siguiente tabla III.

Además, el porcentaje de correlaciones sin contaminar (PUC) es de 0,75. Valor inferior a 0,8, que, en combinación con el ECV de 0,669, pone en duda la presencia única de un factor general.

TABLA III. Índice de varianza común explicada el ítem (I_ECV)

Ítem	I_ECV	Ítem	I_ECV
I1	0,469	I11	0,419
I2	0,455	I12	0,632
I3	0,727	I13	0,471
I4	0,433	I14	0,498
I5	0,172	I15	0,256
I6	0,507	I16	0,469
I7	0,172	I17	0,805
I8	0,780	I18	0,485
I9	0,209	I19	0,193
I10	0,395	I20	0,270

Fuente: Elaboración propia.

En cuarto lugar, el índice H de replicabilidad del constructo señala la capacidad de representación que tienen el conjunto de ítems para definir cada factor. Los resultados se muestran en la siguiente tabla:

Se logra, en todos los casos, valores cercanos o superiores a 0,7. Por tanto, el 70% o más de la variabilidad de cada factor latente está determinado de indicadores que lo componen y puede considerarse aceptable.

TABLA IV. Índice de replicabilidad del constructo (H)

	F. General	Cognitivo	Creativo	Social	Emocional
H	0,869	0,697	0,865	0,678	0,721

Fuente: Elaboración propia.

Finalmente, como muestra la Tabla V, los pesos factoriales muestran cargas moderadas para el factor general, entre 0,3 y 0,6, pero todas significativas. Los valores en las subdimensiones, aunque similares, son algo más altos y también resultan significativas, como muestra la siguiente tabla. En promedio, tanto el factor general como las subdimensiones, tienen cargas factoriales considerables, superando el punto de corte propuesto de 0,3. En el caso del factor general se encuentra muy cerca de

TABLA V. Pesos factoriales estandarizados, R^2 e índice de sesgo relativo de los parámetros (SRP) del modelo bifactorial (M9)

	λ*					λ*		
Ítem	**F. General**	**Cognitiva**	**Creativa**	**Social**	**Emocional**	**R^2**	**Unidim.**	**SRP**
I2	0,602	0,659				0,797	0,696	0,034
I3	0,632	0,387				0,549	0,602	0,156
I4	0,504	0,577				0,587	0,574	0,047
I13	0,595	0,631				0,751	0,692	0,139
I14	0,404	0,406				0,328	0,399	0,066
I7	0,339		0,745			0,670	0,740	0,054
I9	0,407		0,792			0,793	0,829	1,183
I10	0,577		0,714			0,842	0,897	0,169
I11	0,570		0,671			0,775	0,854	1,037
I12	0,438		0,334			0,303	0,508	0,555
I17	0,461		0,227			0,264	0,467	0,498
I18	0,520		0,536			0,558	0,701	0,160
I20	0,432		0,711			0,692	0,773	0,163
I1	0,446			0,475		0,424	0,431	0,012
I6	0,517			0,510		0,527	0,489	0,010
I16	0,548			0,583		0,639	0,527	0,038
I19	0,341			0,698		0,604	0,390	0,013
I5	0,362				0,793	0,761	0,386	0,348
I8	0,544				0,289	0,379	0,452	0,144
I15	0,390				0,665	0,595	0,394	0,789
Promedio	0,481	0,532	0,591	0,567	0,582	0,592	0,590	0,281

*Todos los parámetros son significativos (p<,001).

Recibido: 27/05/2024 Aceptado: 07/06/2024

0,5 y en las subdimensiones algo por encima de ese valor y con niveles muy similares entre ellas.

Observando los valores de los pesos factoriales de la tabla anterior, en 5 ítems de los 20 que componen la escala, el valor es mayor en el factor general que en las subdimensiones. Son el ítem 3, 6, 8, 12 y 17, pertenecientes a las diferentes subdimensiones del modelo. Y el modelo bifactorial (M9) logra explicar aproximadamente el 60% de variabilidad de los datos.

Respecto al sesgo relativo de los parámetros (SRP), el valor promedio está cerca del 30%, por tanto, al exceder el límite del 15%, queda constancia de las diferencias sustanciales en los efectos de ese factor general en los dos modelos.

Conclusiones

Los modelos reflectivos no pueden considerarse superiores a los formativos, o viceversa. Ambas pueden ser alternativas en el estudio de la validez de constructo (Bollen y Diamantopoulos, 2017). No obstante, en el caso concreto de la escala GRS la opción reflectiva utilizada en el estudio de la validez de constructo de Tourón et al. (2023), se ajusta a la definición teórica y operativa del mismo. Considerando los resultados del modelo bifactorial, se confirma la presencia de un factor general como causante de gran parte de la variabilidad de las respuestas a los ítems y también de una estructura multidimensional de factores específicos que aporta otra parte a la varianza común del modelo (capacidad cognitiva, capacidad creativa, habilidades sociales y control emocional).

El valor del índice de unidimensionalidad (ECV) y el porcentaje de correlaciones sin contaminar (PUC), con valores inferiores a 0,8 en ambos casos, no muestra la presencia clara de un único de un factor general para explicar las respuestas a los ítems. No obstante, estos resultados muestran la importancia de ese factor, pero también la influencia de las subdimensiones en la explicación de las diferencias. Y, como señalan los coeficientes omega jerárquicos de las subdimensiones, todos cercanos o superiores a 0,5, pueden considerarse efectos importantes, Smits et al. (2014). Raise et al. (2013) apuntan que con valores de PUC inferiores a 0,7, un valor de ECV del factor general superiores a 0,6 y un omega jerárquico superior a 0,7 sugieren la presencia de cierta

multidimensionalidad, aunque no descartan totalmente la interpretación de la escala como unidimensional.

Las diferencias de intensidad existentes entre los pesos factoriales de cada factor demuestran el carácter congenérico de las medidas, pero los resultados señalan la presencia de una causa común que determina gran parte de la variabilidad de la matriz de varianzas-covarianza. Ese factor general identificado en el modelo factorial determina aproximadamente el 70% (ECV=0,71) de la variabilidad de la varianza común y, por tanto, la intercambiabilidad de los indicadores, considerados conductas provocadas por el mismo constructo, puede mantenerse. Esta cuestión es clave para definir un modelo de medida como reflectivo (Murray y Booth, 2018). El resto de la varianza común, un 30% aproximadamente, está producida por las diferencias en los resultados de las subdimensiones.

Los pesos factoriales han resultados significativos en todos los casos, tanto en el factor general como en las subdimensiones. Y los índices ECV de cada ítem (ver Tabla III) han mostrado que el factor general determina una parte de las respuestas a los ítems, pero solo en el caso del ítem 17 esa aportación supera el 80%. En los ítems 3 y 8 el factor general explica aproximadamente el 75%. En 10 de los ítems, la aportación se sitúa entre el 41%-60% y en los siete ítems restantes es aproximadamente del 20%-40%. Por tanto, en la mayoría de los ítems, la influencia del factor general se combina con el impacto de las subdimensiones.

Además, los índices H de replicabilidad del constructo muestran buenos resultados, con valores cercanos o superiores a 0,7. Por tanto, el conjunto de indicadores que componen cada variable latente explica una variabilidad suficiente y tendrá estabilidad en diferentes estudios (Rodríguez et al., 2016).

Considerando los valores de las cargas factoriales, en el modelo bifactorial muestran un tamaño suficiente tanto en el factor general como en las subdimensiones, con valores promedio cercanos o superiores a 0,5, mostrando la importancia del modelo completo (Ferrando y Lorenzo-Seva, 2018). También el índice de sesgo relativo de los parámetros (SRP) apunta que las cargas factoriales del factor general en el modelo bifactorial son diferentes a las del modelo unidimensional. Por tanto, no considerar la estructura multidimensional produce sesgo en las estimaciones.

Agradecemos toda la revisión crítica llevada a cabo y esperamos haber respondido a la mayor parte de las objeciones planteadas de manera satisfactoria. Tendremos en cuenta las nuevas vías de consideración que se abren como consecuencia de esta discusión y reanálisis de nuestros

Recibido: 27/05/2024 Aceptado: 07/06/2024

datos. Finalmente agradecemos al editor de la revista que haya aceptado incluir estos trabajos a raíz del nuestro original, pues entendemos que estos debates y diferencias de posturas metodológicas hacen progresar y mejorar el trabajo científico.

Referencias bibliográficas

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas (M. Lieve, Trans.). AERA*. American Educational Research Association.

Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bollen, K. A., & Diamantopoulos, A. (2017). In defense of causal-formative indicators: A minority report. *Psychological Methods*, *22*(3), 581–596. https://doi.org/10.1037/met0000056

Bollen, K.A. & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110(2)*, 305–314. https://doi.org/10.1037/0033-2909.110.2.305

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A Comparison of Bifactor and Second-Order Models of Quality of Life. *Multivariate Behavioral Research*, *41*(2), 189–225. https://doi.org/10.1207/s15327906mbr4102_5

Chen, F.F., & Zhang, Z., (2018). Bifactor Models in Psychometric Test Development. En P. Irwing, T. Booth y D.J. Hughes, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test*, 325-345. https://doi.org/10.1002/9781118489772.ch12

Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis. *Educational and psychological measurement*, 78(5), 762-780. https://doi.org/10.1177/0013164417719308

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller

(Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). IAP Information Age Publishing.

Hayduk, L. A. (2014a). Seeing perfectly-fitting factor models that are causally misspecified: Understanding that close-fitting models can be worse. *Educational and Psychological Measuremen*t, *74*(6), 905-926. https://doi.org/10.1177/0013164414527449

Hayduk, L. A. (2014b). Shame for disrespecting evidence: The personal consequences of insufficient respect for structural equation model testing. *BMC: Medical Research Methodology, 14,* 124 https://doi.org/10.1186/1471-2288-14-124

Holzinger, K.J., and Swineford, F. (1937). The Bi-factor method. *Psychometrika 2,* 41–54. https://doi.org/10.1007/BF02287965

Hu, L. y Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6(1),* 1-55. https://doi.org/10.1080/10705519909540118

Jabůrek, M., Ťápal, A., Portešová, Š., Pfeiffer, S. I. (2020). Validity and Reliability of Gifted Rating Scales-School Form in Sample of Teachers and Parents – A Czech Contribution. *Journal of Psychoeducational Assessment*, *39*(3), 361–371. https://doi.org/10.1177/0734282920970718

Jöreskog, K. G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. *Quality and Quantity*, *24*, 387–404.

Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika*, *59*, 381–389.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.

McDonald, R.P. (1999). *Test Theory: A Unified Treatment (1st ed.).* Psychology Press. https://doi.org/10.4324/9781410601087

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*(3), 568-592. https://doi.org/10.1037/0021-9010.93.3.568

Murray, A.L. & Booth, T. (2018). Causal Indicators in Psychometrics. InP. Irwing, T. Booth and D.J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing*. https://doi.org/10.1002/9781118489772.ch7

Muthén, B. (1984). A general structural model with dichotomous, ordered categorical and continuous latent variable indicators. *Psychometrika*, *49*(1), 115–132. https://doi.org/10.1007/BF02294210

Pfeiffer, S. (2017). *Identificación y evaluación del alumnado con altas capacidades: Una guía práctica*. La Rioja: UNIR Editorial.

Pfeiffer, S. I., & Jarosewich, T. (2003). *GRS: Gifted Rating Scales*. Psychological Corporation.

Pfeiffer, S. I., Petscher, Y., & Kumtepe, A. (2008). The Gifted Rating Scales-School Form: A Validation Study Based on Age, Gender, and Race. *Roeper review, 30*(2), 140–146. https://doi.org/10.1080/02783190801955418

Pfeiffer, S. I., Shaunessy-Dedrick, E., & Foley-Nicpon, M. (Eds.). (2018). *APA handbook of giftedness and talent. American Psychological Association*. https://doi.org/10.1037/0000038-000

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. http://dx.doi.org/10.1080/00273171.2012.715555

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and Structural Coefficient Bias in Structural Equation Modeling: A Bifactor Perspective. *Educational and Psychological Measurement*, *73*(1), 5-26. https://doi.org/10.1177/0013164412449831

Renzulli, J. S., & Reis, S. M. (2018). The three-ring conception of giftedness: A developmental approach for promoting creative productivity in young people. In S. I. Pfeiffer, E. Shaunessy-Dedrick, & M. Foley-Nicpon (Eds.), *APA Handbook of Giftedness and Talent* (pp. 185–199). American Psychological Association. https://doi.org/10.1037/0000038-012

Rodríguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137–150. https://doi.org/10.1037/met0000045

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online, 8*(2), 23-74.

Shi, D., DiStefano, C., McDaniel, H. L., & Jiang, Z. (2018). Examining Chi-Square Test Statistics Under Conditions of Large Model Size and Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(6), 924-945. https://doi.org/10.1080/10705511.2018.1449653

Smits, I. A., Timmerman, M. E., Barelds, D. P., & Meijer, R. R. (2014). The Dutch symptom checklist-90-revised. *European Journal of Psychological Assessment*, *31*(4), 263-271. https://doi.org/10.1027/1015-5759/a000233

Stucky, B. D., & Edelen, M. O. (2014). Using hierarchical IRT models to create unidimensional measures from multidimensional data. En S. P. Reise & D. A. Revicki (Eds.), *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment* (pp. 183-206). Routledge/Taylor & Francis.

Tourón, J. (2012). *¿Superdotación o alta capacidad?* Recuperado de https://www.javiertouron.es/superdotacion-o-alta-capacidad/

Tourón, J. (2020). Las altas capacidades en el sistema educativo español: reflexiones sobre el concepto y la identificación: Concept and Identification Issues. *Revista de Investigación Educativa*, *38*(1), 15–32. https://doi.org/10.6018/rie.396781

Tourón, J. (2023). ¿Puede una escuela inclusiva ignorar a sus estudiantes con altas capacidades? *enTERA2.0, 10*, 24-46.

Tourón, M., Navarro-Asensio, E., Tourón, J. (2023). Validez de Constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres, (GRS 2), en España. *Revista de Educación*, *402*, 55-83. https://doi.org/10.4438/1988-592X-RE-2023-402-595

Tourón, M., Tourón, J., & Navarro-Asencio, E. (2024). Validación española de la Escala de Detección de altas capacidades, «Gifted Rating Scales 2 (GRS 2-S) School Form», para profesores. *Estudios Sobre Educación*, *46*, 33-55. https://doi.org/10.15581/004.46.002

Wang, J. y Wang, X. (2020). *Structural Equation Modeling: Applications Using Mplus, Second Edition*. John Wiley & Sons Ltd.

Wu, M., Tam, H. P., & Jen, T. H. (2016). *Educational measurement for applied researchers. Theory into Practice*. https://doi.org/10.1007/978-981-10-3302-5

Zumbo, B. D. (2006). 3 validity: foundational issues and statistical methodology. En C. R. Rao, & S. Sinharay (Eds.), *Handbook of Statistics: Vol. 26. Psychometrics, 45-79*. Elsevier Science. https://doi.org/10.1016/S0169-7161(06)26003-6

Información de contacto: Marta Tourón. Universidad Internacional de La Rioja - UNIR. Avd. de La Paz, 137. 26006 Logroño La Rioja. E-mail: marta.tporto@unir.net

Recibido: 27/05/2024 Aceptado: 07/06/2024

Contribución al conocimiento en Investigación Educativa a partir de estudios de validación de escalas: una reflexión crítica sobre el trabajo de M. Tourón et al. (2023) y el análisis metodológico de Martínez-García (2024)

Contribution to knowledge in Educational Research through scale validation studies: a critical review on the work of M. Tourón et al. (2023) and the methodological analysis by Martínez-García (2024)

https://doi.org/10.4438/1988-592X-RE-2024-406-637

Fernando Martínez Abad
https://orcid.org/0000-0002-1783-8198
Universidad de Salamanca

José Carlos Sánchez Prieto
https://orcid.org/0000-0002-8917-9814
Universidad de Salamanca

Resumen

Existe en la literatura científica un gran volumen de investigaciones que proponen el diseño y validación de escalas de medida en el ámbito de las ciencias de la educación. Este trabajo presenta, desde el punto de vista de la investigación educativa aplicada, un análisis crítico de la validación de la escala de detección de alumnos con altas capacidades (GRS 2) para padres publicada por M. Tourón et al. (2023) en Revista de Educación, teniendo en cuenta la crítica metodológica realizada por Martínez-García (2024) a dicha validación. La presente propuesta pretende sumarse a este diálogo mediante el análisis de ambas publicaciones y la formulación de cuestiones de mejora, centrándose

Recibido: 01/06/2024 Aceptado: 01/06/2024

especialmente en la fundamentación teórica de los modelos, la formulación de compuestos reflectivos y formativos, y la adecuación del uso de indicadores de bondad de ajuste alternativos al Chi Cuadrado en el modelo de análisis factorial confirmatorio. Se observa la aplicación de ciertos procedimientos muy asentados en los estudios de validación psicométrica dentro de la investigación educativa que, además de no promover la comprensión de las bases teóricas en las que se fundamenta la escala, dificultan su adecuado empleo posterior por parte de investigadores aplicados en estudios diagnósticos o experimentales. Así, los argumentos analizados ponen de relieve la necesidad de replantear, a nuestro juicio, algunas de las prácticas habituales en este tipo de estudios que disminuyen su relevancia y las implicaciones de los resultados obtenidos para el desarrollo de la teoría y la práctica educativa.

Palabras clave: validación de escalas, análisis factorial exploratorio, análisis factorial confirmatorio, modelos de ecuaciones estructurales, modelos formativos.

Abstract

In the scientific literature, there is a considerable volume of research proposing the design and validation of measurement scales in the field of educational sciences. This work presents, from the perspective of applied educational research, a critical analysis of the validation of the detection scale for high-ability students (GRS 2) for parents published by M. Tourón et al. (2023) in Revista de Educación, taking into account the methodological critique made by Martínez-García (2024) of said validation. The present proposal aims to contribute to this dialogue by analyzing both publications and formulating improvement issues, focusing especially on the theoretical foundation of the models, the formulation of reflective and formative compounds, and the adequacy of using alternative goodness-of-fit indicators to Chi Square in the confirmatory factor analysis model. Certain procedures deeply established in psychometric validation studies within educational research are observed, which, besides not promoting the understanding of the theoretical foundations underlying the scale, hinder its subsequent appropriate use by applied researchers in diagnostic or experimental studies. Thus, the analyzed arguments highlight the need to reconsider, in our view, some of the common practices in this type of studies that diminish their relevance and the implications of the results obtained for the development of educational theory and practice.

Keywords: scales validation, explanatory factor analysis, confirmatory factor analysis, structural equations modelling, formative models.

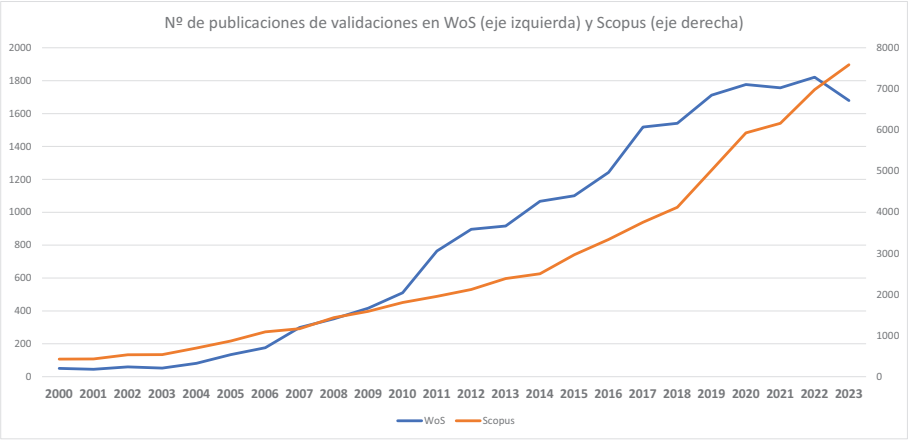
Introducción

Los estudios psicométricos orientados al diseño y validación de escalas e instrumentos conforman una línea de investigación muy popular y profusa en Investigación Educativa. Esto se refleja en una amplia presencia en publicaciones científicas del área de Ciencias Sociales y de la Educación, como ilustra la Figura I. De hecho, el aumento del interés por el diseño y validación de escalas de medida es exponencial, con un importante crecimiento en el número de publicaciones científicas desde 2005 en ambas bases de datos.

Como evidencia de esta amplia presencia, cada vez son más las revistas que argumentan esta casuística como justificación en el rechazo de propuestas, llegando a estar recogido en la normativa para los autores en algunas de ellas (e.g., Educación XX1, 2024).

Un buen ejemplo de este tipo de trabajos es el estudio de M. Tourón et al. (2023), publicado recientemente en la Revista de Educación. El estudio aborda una validación psicométrica en la población española de la Escala de Detección de alumnos con Altas Capacidades para Padres

FIGURA I. Evolución del número de publicaciones de validaciones de escalas en WoS y Scopus



Nota: Búsqueda en WoS en fuentes del Área *Education & Educational Research*, y en Scopus en fuentes del Área *Social Sciences*. Cadena de búsqueda: *(validation OR design OR psychometric) AND (scale OR instrument)*.

(GRS 2), incluyendo una exploración inicial mediante un análisis factorial exploratorio (AFE) y, a partir de estos resultados previos, la validación de ocho modelos dimensionales diferentes propuestos mediante un análisis factorial confirmatorio (AFC), hasta la selección del modelo más adecuado a los datos.

El proceso de diseño y validación estadística de escalas es potencialmente valioso para la comunidad científica, ya que posibilita la medida adecuada de rasgos o constructos a una población objetivo en estudios de carácter aplicado, ya sean exploratorios, correlacionales o experimentales. Así, bajo nuestro punto de vista, la presentación de los resultados de estudios psicométricos debe documentar clara y sistemáticamente toda la información que permita a otros investigadores educativos aprovechar y utilizar su potencial.

En este contexto, y teniendo en cuenta que el análisis crítico de Martínez-García (2024) se centra fundamentalmente en las decisiones metodológicas tomadas por M. Tourón et al. (2023) en el proceso de validación estadística, la presente revisión pretende contribuir al debate poniendo el foco principal en la aportación práctica que supone para los investigadores educativos aplicados tanto la validación de la escala GRS 2 propuesta por M. Tourón et al. (2023) como las críticas de Martínez-García (2024).

Este documento se divide en cuatro apartados principales. En primer lugar, se presenta una reflexión en torno a la aportación que realiza el trabajo de M. Tourón et al. (2023), junto con las críticas de Martínez-García (2024), al desarrollo científico en el ámbito de las Ciencias de la Investigación. Posteriormente, se analiza la pertinencia de las decisiones metodológicas tomadas en el artículo de validación psicométrica, poniendo el foco en si estas decisiones contribuyen a una mejor comprensión de la escala validada, fomentando el empleo de la misma en estudios aplicados y, finalmente, un avance en el conocimiento sobre el estudio de las altas capacidades. El tercer bloque de este trabajo se centra en desgranar y discutir las principales críticas de Martínez-García (2024), de nuevo bajo un punto de vista más práctico-aplicado que a partir del enfoque teórico-básico que desarrolla el autor del análisis crítico. Finalmente, se incluye un bloque a modo de conclusión en el que se reflexiona en torno a lo generalizado en la investigación educativa de las dudosas prácticas aquí identificadas, y la necesidad de una toma de conciencia sobre las mismas que lleve al desarrollo de trabajos con una orientación que no se centre principalmente en la publicabilidad.

Recibido: 01/06/2024 Aceptado: 01/06/2024

Contribución al corpus científico en investigación educativa

Cabe destacar inicialmente que valoramos muy positivamente los procedimientos técnicos y decisiones estadísticas implementadas en el trabajo de M. Tourón et al. (2023). Tal y como destaca Martínez-García (2024), el estudio de validación realiza un esfuerzo importante por desarrollar un procedimiento sistemático y conforme a los estándares y usos más habituales en la investigación psicométrica. No obstante, encontramos algún obstáculo importante que dificulta la trazabilidad y aplicabilidad de los resultados obtenidos y presentados.

En primer lugar, a pesar de que la escala GRS 2 está publicada originalmente en lengua inglesa, y que los ítems fueron traducidos para llevar a cabo el proceso de validación, llama la atención que M. Tourón et al. (2023) no presentan el listado completo de los ítems traducidos. Esta omisión dificulta la adecuada comprensión y seguimiento de las decisiones tomadas en relación a los modelos dimensionales probados en el trabajo, a la vez que imposibilita el empleo de la escala en estudios aplicados. Dado que la escala original pertenece a una editorial (MHS), lo más probable es que no haya sido posible publicar la escala completa. Esta cuestión la indican los autores de manera indirecta cuando se refieren a la traducción de los ítems. Cuando menos, este tipo de prácticas no son las más adecuadas para contribuir al avance del conocimiento científico, la ciencia abierta y el desarrollo de futuros trabajos académicos.

Por otro lado, a pesar del esfuerzo realizado por M. Tourón et al. (2023) en definir todos los modelos dimensionales plausibles a partir del conjunto de ítems y probar su ajuste, los autores no dedican ese mismo esfuerzo en justificar la pertinencia y sentido teórico de cada uno de los modelos validados.

Si bien es cierto que los autores incluyen en las conclusiones del trabajo una reflexión teórica en torno al modelo finalmente seleccionado, la omisión de una justificación más amplia sobre todos los modelos probados no sólo dificulta al lector una mejor comprensión de la escala GRS 2 y de su estructura dimensional, sino que limita sensiblemente el alcance del estudio al reducir el análisis de la relación entre teoría y praxis.

A este respecto, habría sido de utilidad para la comunidad científica una mayor discusión sobre las implicaciones de los resultados para la conceptualización de la inteligencia y la relación entre sus distintos componentes, resultando especialmente llamativa la división de las tres

Recibido: 01/06/2024 Aceptado: 01/06/2024

dimensiones iniciales en cuatro durante el AFE a las que se le añadieron dos de segundo orden durante el AFC.

Como añadido, no está de más recordar en este punto los riesgos asociados a la definición de la estructura dimensional de una escala teniendo en cuenta exclusivamente los datos empíricos (Brown, 2015; MacCallum et al., 1992; Schmitt et al., 2018). En este sentido, consideramos que puede existir un sesgo por sobreajuste en los resultados del AFC.

En suma, el trabajo de M. Tourón et al. (2023) no incluye el contenido de los ítems de la escala GRS 2 ni una justificación clara de los 8 modelos dimensionales validados en el AFC aplicado. Estas dos carencias dificultan la labor hermenéutica del artículo e impiden su reproducibilidad al quedar incompleto el proceso de modelado de las dimensiones teóricas de la escala. Con un buen desarrollo teórico, el artículo podría haber ido más allá de la presentación de una herramienta valida y fiable para la detección de altas capacidades, aportando información de interés para la comprensión de los componentes de la inteligencia.

En conjunto, estas limitaciones suponen una barrera para que otros investigadores y profesionales del ámbito educativo puedan seguir avanzando en la detección y estudio de alumnado con altas capacidades.

Procedimiento y decisiones metodológicas

Tal y como hemos indicado anteriormente, tras realizar una exploración inicial de la estructura dimensional mediante AFE, M. Tourón et al. (2023) proponen y comprueban el ajuste de 8 modelos dimensionales diferentes en el AFC. Aunque se sigue el procedimiento habitual en este tipo de análisis, en la redacción del trabajo no queda clara la conexión entre los resultados del AFE y los modelos propuestos en el AFC. En este punto estamos de acuerdo parcialmente con la crítica de Martínez-García (2024) sobre el uso del AFE en el trabajo, al no quedar claro qué aporta el AFE aplicado a las decisiones tomadas en la aplicación del AFC. En este sentido, consideramos que la aplicación del AFE en este artículo es contingente: existe una teoría previa sólida sobre la estructura factorial de la escala, y las evidencias obtenidas en el AFE no conducen a la toma de decisiones sobre qué modelos testear en el AFC.

Aunque luego abordaremos esta cuestión en mayor profundidad, debemos decir que no estamos de acuerdo con las críticas de Martínez-

Recibido: 01/06/2024 Aceptado: 01/06/2024

García (2024) sobre el empleo de índices de ajuste complementarios al contraste Chi Cuadrado. Al igual que los estadísticos de tamaño del efecto son tremendamente valiosos en la adecuada interpretación de los contrastes de hipótesis bivariados (Cohen, 1988; Grissom & Kim, 2011), los indicadores de ajuste absoluto e incremental complementarios a Chi Cuadrado permiten realizar un análisis más fino y claro del grado de similitud entre el modelo empírico y los modelos nulo y saturado.

Como se ha mencionado anteriormente, una de las limitaciones principales de este tipo de estudios es el carácter puramente empirista de la toma de decisiones. Al respecto de los modelos probados en el AFC, los autores incluyen modelos con errores correlacionados tomando en cuenta la información aportada por los índices de modificación. Esta práctica no es la más recomendable, ya que puede dar lugar a modelos artificialmente sobreajustados y difícilmente interpretables en la práctica (MacCallum et al., 1992; Schmitt et al., 2018). Teniendo en cuenta que la escala original está previamente validada en una población similar, otra decisión que genera dudas es el cambio de dimensión que sufre el ítem 17 en alguno de los modelos. Dado que los autores no informan sobre los pesos factoriales de los modelos AFE aplicados, y no se conoce la redacción de los ítems, no se aportan evidencias suficientes que justifiquen esta decisión.

Una crítica fundamental de Martínez-García (2024) es la relativa a la decisión no justificada de M. Tourón et al. (2023) de considerar las dimensiones de la escala GRS 2 como reflectivas. Los autores no plantean la posibilidad de que los ítems, o el constructo de segundo orden probado en alguno de los modelos, puedan conformar una estructura formativa. Debemos tener en cuenta que la presencia de estructuras dimensionales formativas en los estudios psicométricos del ámbito de las ciencias de la educación es residual. En este sentido, consideramos importante dar a conocer en el campo de la investigación educativa las características y propiedades de los factores formativos, en relación a la naturaleza y supuestos de los factores reflectivos.

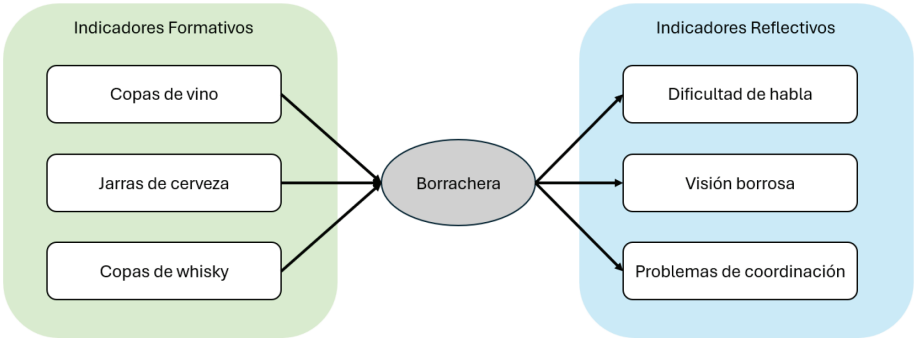
Un factor reflectivo es aquel en el que las variables latentes se consideran compuestos en los que los indicadores son el resultado de dicha variable, de esta manera, un cambio en el compuesto reflectivo provoca un cambio en todos los indicadores. Por otro lado, en el caso de los factores formativos las variables latentes son consideradas compuestos en los que los indicadores son su causa de forma que un cambio en el

compuesto no va necesariamente acompañado de un cambio en todos los indicadores (Starsted et al., 2016). Esta diferencia en la relación entre compuesto e indicador se representa gráficamente mediante la dirección de las flechas que unen el compuesto con sus indicadores, mientras que en los compuestos reflectivos las flechas van del compuesto hacia sus indicadores (al igual que sucede en los modelos CB-SEM), en el caso de los compuestos formativos las flechas van en la dirección contraria.

Un ejemplo clásico de la diferencia entre compuestos reflectivos y formativos es el *ejemplo del borracho* (Figura II) (Chin, 1998). En este caso, si modelásemos la borrachera de manera reflectiva este tendría indicadores como "visión doble", "dificultad en el habla" o "problemas de coordinación" de manera que un incremento en la borrachera generaría un aumento en sus indicadores. De manera opuesta, si la borrachera fuese un compuesto formativo esta incluiría indicadores como "número de cervezas", "número de copas de vino" o "número de copas de whisky", en este caso un aumento de la borrachera no implica un aumento en todos los indicadores.

La cuestión del uso de modelos que incluyen compuestos formativos constituye un punto abierto de debate, que se entremezcla con la disputa en la comunidad científica sobre la validez del análisis confirmatorio de compuestos (ACC) que sólo puede llevarse a cabo mediante la aplicación de modelos de ecuaciones estructurales basados en mínimos cuadrados parciales (PLS-SEM), una técnica de naturaleza predictivo-causal (Hair et al., 2011). Esta cuestión tiene implicaciones que se abordarán más adelante.

FIGURA II. Diferencia entre indicadores reflectivos y formativos



Fuente: Elaboración propia.

Recibido: 01/06/2024 Aceptado: 01/06/2024

En lo que se refiere a la crítica realizada por Martínez-García (2024), resulta difícil precisar si la escala validada por los autores incluye compuestos reflectivos o formativos, dado que no se aborda esta cuestión en el artículo ni, como ya se ha mencionado, se incluye el listado completo de los ítems que la componen para que el lector pueda juzgar esta cuestión por sí mismo. No obstante, en nuestra opinión, a juzgar por alguno de los ítems mencionados y teniendo en cuenta las estructuras teóricas planteadas, existe la posibilidad de que se trate de compuestos formativos al menos en alguno de los casos. En resumen, aunque no estamos de acuerdo con todas las críticas realizadas por Martínez-García (2024), sí que se puede considerar que el artículo de M. Tourón et al. (2023) presenta algunas limitaciones derivadas de la falta de justificación teórica de los modelos propuestos, la toma de decisiones basadas en datos que pueden provocar sesgos por sobreajuste y el uso de factores reflectivos sin una consideración previa. Estos problemas no son exclusivos del artículo objeto de debate, sino que ejemplifican cuestiones que se encuentran ampliamente en la literatura, producto del proceso de *escritura automática* de artículos (según las palabras del propio Martínez-García (2024)) en el que muchas veces se cae debido a las presiones para intensificar la labor de difusión de resultados (Martínez-García & Martínez-Caro, 2009).

Análisis crítico de Martínez-García (2024)

Como señalábamos anteriormente, uno de los pilares de la crítica de Martínez-García (2024) al trabajo de M. Tourón et al. (2023) es la definición predeterminada de los factores latentes como reflectivos en todos los modelos definidos, sin contemplar la posibilidad de que sean formativos. Si bien estamos de acuerdo en el fondo con esta crítica, es importante señalar que la argumentación de Martínez-García (2024) se basa exclusivamente en cuestiones estadísticas y matemáticas, sin evidenciar la naturaleza formativa de estos factores en base a la redacción de sus ítems. De hecho, consideramos que Martínez-García (2024) presenta una visión reduccionista de los modelos reflectivos, cuando indica:

Los indicadores observables son reflectivos […] por lo que con un solo indicador se podría medir cada una de esas dimensiones. Si un investigador piensa que con un solo indicador es insuficiente para medir

una variable latente y que necesita más porque la dimensión latente es amplia […], lo que probablemente esté sucediendo es que esa batería de indicadores está midiendo variables latentes diferentes (Martínez-García, 2024)

Estamos en profundo desacuerdo con esta afirmación, bajo nuestro punto de vista existen fenómenos que presentan un amplio abanico de respuestas o conductas medibles que se espera que estén correlacionadas entre sí, y que en conjunto aportan matices, riqueza y una mayor exactitud en la medida del factor. Como ilustración de esta cuestión, solamente hace falta recordar el anterior *ejemplo del borracho*, en el que las posibles manifestaciones de la borrachera son muy variadas, por lo que, aunque estén correlacionadas, puede considerarse que si no se incluyeran todas las manifestaciones posibles se estaría reduciendo la precisión en la medida del factor.

Otra de las cuestiones clave que señala Martínez-García (2024), que ya hemos abordado brevemente más arriba, pero en la que nos queríamos detener más específicamente, es la referida al empleo de índices de ajuste absoluto e incremental complementarios al contraste chi-cuadrado. Martínez-García (2024) es categórico al afirmar que estos indicadores "comúnmente son usados para ajustar los modelos de forma que se evite el único test adecuado para detectar la mala especificación de un modelo (el test de la chi-cuadrado o, más correctamente, la familia de test de la chi-cuadrado)". A lo que añade que "cuando un modelo no pasa el test de la chi-cuadrado, los parámetros estimados no son interpretables, ya que una o varias de las relaciones de covarianza implicadas por el modelo no es apoyada por los datos empíricos, lo que de inmediato produce sesgo en las estimaciones" (Martínez-García, 2024). Podríamos aceptar este posicionamiento si el test chi-cuadrado fuera infalible, pero es bien conocido que esta prueba devuelve puntuaciones hinchadas (lo que supone una muy elevada tasa de rechazo de los modelos) ante condiciones de falta de normalidad multivariante (e.g., Curran et al., 1996; Hu et al., 1992), o que, como el propio Martínez-García (2024) admite, puede fallar en la detección de modelos mal especificados. Además, Martínez-García (2024) contrapone en su discusión la amplia literatura existente de autores de referencia en este campo, que "critican la prueba chi-cuadrado y recomiendan los índices de ajuste aproximados", con la crítica que realiza a estos índices Hayduk (2014), un investigador con niveles de difusión e impacto relativos en el ámbito de la investigación

Recibido: 01/06/2024 Aceptado: 01/06/2024

básica sobre modelización con estructuras de covarianzas claramente inferiores[1]. En este punto debemos preguntarnos si la relativa irrelevancia de Hayduk en comparación con investigadores de referencia como Jöreskog, Bentler, Steiger o Browne, se debe a que su trabajo no realiza una aportación significativa al conocimiento o a si su discurso resulta inconveniente para la publicabilidad de estudios que incluyen modelos de ecuaciones estructurales (SEM) y para la explotación comercial de los principales paquetes estadísticos que permiten su obtención.

Por último, al examinar en conjunto las críticas de Martínez-García (2024) apreciamos que no repara en la problemática que supone el estudio de los niveles de ajuste de los modelos con compuestos formativos. Es importante señalar que, como ya se ha mencionado, la definición de compuestos formativos no es formalmente posible en los modelos tradicionales basados en estructuras de covarianzas, o MEE. La teorización y validación de factores formativos es viable a través de métodos de mínimos cuadrados parciales (PLS-SEM), basados en las estructuras de varianzas de los datos. A diferencia de las técnicas basadas en estructuras de covarianzas, el PLS-SEM es más flexible en cuanto a los supuestos previos y el ajuste a distribuciones teóricas, siendo considerada una técnica de carácter exploratorio y confirmatorio que estudia relaciones de tipo predictivo-causal (Hair et al., 2022). De hecho, no es posible el empleo del test chi cuadrado para la verificación del del ajuste de los modelos PLS-SEM. Lo que es más, a pesar de que las herramientas actuales de PLS-SEM permiten estudiar el ajuste de los modelos a través de indicadores de ajuste aproximados (e.g., RMSEA, CFI, TLI), su empleo no está recomendado por buena parte de los investigadores de referencia en el ámbito (Hair et al., 2022). A esto se le suma, que en el caso de los compuestos formativos tampoco se aplican otros supuestos previos presentes en CFA como el análisis de las cargas de los ítems, el índice de fiabilidad compuesta o la varianza media extraída.

En suma, a pesar de que el análisis crítico de Martínez-García (2024) preconiza del uso del test chi cuadrado como única evidencia válida para la verificación del buen o mal ajuste del modelo, a la vez que defiende

[1] El trabajo fundamental de Hayduk en el que Martínez-García fundamenta su crítica (Hayduk, 2014) alcanza, en febrero de 2024, un total de 55 citas en Google Scholar, 40 citas en Web of Science, y 41 citas en Scopus.

Recibido: 01/06/2024 Aceptado: 01/06/2024

la existencia de dimensiones formativas en el modelo factorial de la escala GRS-2, no propone soluciones sobre cómo validar el ajuste de este modelo con dimensiones formativas sin el empleo de indicadores de ajuste aproximados.

Conclusiones

Podemos afirmar que el trabajo de M. Tourón et al. (2023), con sus fortalezas y sus limitaciones, es un buen ejemplo de cómo se están diseñando e implementando los procesos de validación estadística de escalas en Investigación Educativa. Así, las problemáticas abordadas en nuestro análisis crítico no son específicas de la propuesta de M. Tourón et al. (2023) y de Martínez-García (2024), sino que son cuestiones observadas de manera generalizada en la investigación psicométrica del ámbito de las Ciencias de la Educación.

En primer lugar, es habitual encontrar trabajos de validación psicométrica que no documentan la redacción completa de los ítems que conforman las escalas validadas (e.g., Fumero & Miguel, 2023; Quijada et al., 2020). Una de las razones fundamentales, como parece ocurrir en M. Tourón et al. (2023), es la existencia de derechos de propiedad intelectual sobre las escalas validadas. En otras ocasiones, nos encontramos con estudios publicados en un idioma diferente al de las escalas aplicadas en la muestra piloto. En estos casos es habitual que los ítems de las escalas se publiquen traducidas al idioma de la revista y no en el idioma original (e.g., Hernández Ramos et al., 2014; Quijada et al., 2020; Sánchez-Prieto et al., 2019), creando sesgos potenciales y dificultando el correcto uso de las mismas. A pesar de que estas prácticas son poco recomendables, están tan asentadas en el ámbito de las Ciencias Sociales y de la Educación que en muchos casos no son detectadas por editores, revisores o lectores especializados en la materia y son asumidas por los propios científicos como una buena praxis. Esta cuestión queda perfectamente ilustrada en las críticas de Martínez-García (2024), que más allá de no mencionar esta omisión, pone encima de la mesa la naturaleza formativa de las dimensiones sin conocer la lectura de sus ítems.

Otra cuestión importante es el desarrollo de un enfoque excesivamente empirista en los procesos de validación estadística de escalas, obviando en gran medida el modelo teórico desde el que ha sido construida la escala

Recibido: 01/06/2024 Aceptado: 01/06/2024

original. Esta perspectiva basada fundamentalmente en la búsqueda del modelo empírico con mejor ajuste puede llevar a soluciones con poco sentido teórico y práctico. Un buen ejemplo de este tipo de prácticas son trabajos que comparan el ajuste de varios modelos empíricamente plausibles sin justificar su relevancia teórica (e.g., Rodríguez Conde et al., 2012; Thomas et al., 2019), estudios que liberan parámetros de correlación entre los errores de algunos ítems basándose exclusivamente en los índices de modificación obtenidos (e.g., Thomas et al., 2019; J. Tourón et al., 2018), o investigaciones que intercambian los ítems entre las dimensiones disponibles, o incluso crean dimensiones nuevas, sin tener en cuenta los modelos teóricos y de indicadores a partir de los que se diseñó la escala (e.g., J. Tourón et al., 2018). Teniendo en cuenta que M. Tourón et al. (2023) no reparan en esta problemática en su análisis de limitaciones, y que Martínez-García (2024) centra sus críticas y realiza propuestas únicamente en relación a la identificación y ajuste de los modelos comparados, nuestra hipótesis es que los investigadores educativos hemos aprehendido y asumido este proceder. Así, nos preguntamos si es necesario reorientar los estudios de validación estadística en investigación educativa: ¿deberíamos sacrificar bondad de ajuste en los modelos (con la consecuente reducción de la publicabilidad de nuestros trabajos) a cambio de que éstos sean más coherentes y explicativos de la realidad? A nuestro juicio, debemos fomentar esta práctica integrando en nuestros procesos de diseño y validación de escalas equipos multidisciplinares que incluyan tanto expertos en los constructos abordados como especialistas en psicometría.

Aunque no profundiza en ello, Martínez-García (2024) destaca la *escritura automática* como una problemática muy existente en Ciencias Sociales. Tanto la falta de un conocimiento estadístico profundo, como asumir una postura acrítica sobre las implicaciones de los procedimientos metodológicos implementados, puede llevar al investigador educativo a la búsqueda de 'recetas' y la aplicación de protocolos de actuación rígidos en la validación de escalas. Este proceder orienta hacia un enfoque puramente empirista, alejado de la necesaria reflexión sobre las implicaciones teóricas y prácticas de los modelos estudiados. Bajo nuestro punto de vista, la aportación efectiva que realizan estos trabajos al avance y la mejora educativa es limitada.

En relación a la escritura automática, y debido fundamentalmente a la presión a la que se ven sometidos los académicos para aumentar su

producción científica, se observa en los últimos años la generalización de prácticas de *salami slicing* (Norman & Griffiths, 2008; Šupak Smolčić, 2013). Esta práctica éticamente dudosa consiste en separar, o trocear, una investigación con sentido unitario en varios informes de resultados, de modo que sea posible la publicación de varios trabajos diferentes. Lamentablemente, el *salami slicing* ha colonizado todas las áreas de la investigación científica, y estudios psicométricos desarrollados en investigación educativa como el aquí analizado no son una excepción (M. Tourón et al., 2023, 2024).

Por último, consideramos fundamental resaltar la falta de conciencia entre los investigadores educativos sobre la existencia de factores formativos, en comparación con los bien conocidos factores reflectivos. Esta falta de conocimiento lleva a que, en la mayor parte de los casos, se establezca un modelo reflectivo por defecto para la validación de escalas sin una reflexión previa sobre el modelo más apropiado. Dados los problemas asociados a tratar de validar factores formativos como reflectivos (Hair et al., 2011), los expertos en la materia debemos redoblar esfuerzos en difundir este tipo de modelos y democratizar su empleo en la investigación psicométrica.

En conclusión, el debate abierto en torno al artículo de M. Tourón et al. (2023) y la posterior crítica de Martínez-García (2024) ofrece la posibilidad de analizar una serie de prácticas naturalizadas en el desarrollo de estudios psicométricos en el ámbito educativo que limitan el desarrollo del campo. A nuestro entender, la sobreabundancia de estudios de esta naturaleza hace necesario dar un salto de calidad para que aumente su relevancia e implicaciones. Este salto de calidad pasa por ampliar el foco de las investigaciones que en la actualidad están excesivamente centradas en la reproducción casi mecanizada de una serie de procesos metodológicos y el cumplimiento de índices de ajuste dejando en segundo plano las implicaciones teóricas y prácticas de los modelos ajustados, que debería ser el objetivo principal de este tipo de estudios.

Referencias bibliográficas

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (Second edition). Guilford Publications.

Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), vii–xvi.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2.ª ed.). Routledge. https://doi.org/10.4324/9780203771587

Curran, P. J., West, S. G, & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16-29.

Educación XX1 (2024). *Normas para la presentación de originales*. https://revistas.uned.es/index.php/educacionXX1/libraryFiles/downloadPublic/170

Fumero, A., & Miguel, A. de. (2023). Validación de la versión española del NEO-FFI-30. *Análisis y Modificación de Conducta*, *49*(179). https://doi.org/10.33776/amc.v49i179.7325

Grissom, R. J., & Kim, J. J. (2011). *Effect sizes for research: Univariate and multivariate applications*. Routledge Academic.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3ª Ed.). Sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hayduk, L. A. (2014). Shame for disrespecting evidence: The personal consequences of insufficient respect for structural equation model testing. *BMC: Medical Research Methodology, 14*(124). https://doi.org/10.1186/1471-2288-14-124

Hernández Ramos, J. P., Martínez Abad, F., García Peñalvo, F. J., Herrera García, M. E., & Rodríguez Conde, M. J. (2014). Teachers' attitude regarding the use of ICT. A factor reliability and validity study. *Computers in Human Behavior*, *31*, 509-516. https://doi.org/10.1016/j.chb.2013.04.039

Hu, L., Bentler, P.M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*(2), 351-362.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490-504. https://doi.org/10.1037/0033-2909.111.3.490

Martínez-García, J.A. (2024, en prensa). Crítica del análisis de la validez de constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres (GRS 2); réplica a Tourón et al. (2023). *Revista de Educación*.

Recibido: 01/06/2024 Aceptado: 01/06/2024

Martínez-García, J. A., & Martínez-Caro, L. (2009). El análisis factorial confirmatorio y la validez de escalas en modelos causales. *Anales de Psicología*, *25*(2), Article 2.

Norman, I., & Griffiths, P. (2008). Duplicate publication and «salami slicing»: Ethical issues and practical solutions. *International Journal of Nursing Studies*, *45*(9), 1257-1260. https://doi.org/10.1016/j.ijnurstu.2008.07.003

Quijada, A., Ruiz, M. A., Huertas, J. A., & Alonso-Tapia, J. (2020). Desarrollo y validación del Cuestionario de Clima Escolar para Profesores de Secundaria y Bachillerato (CES-PSB). *Anales de Psicología*, *36*(1), 155-165. https://doi.org/10.6018/analesps.36.1.341001

Rodríguez Conde, M. J., Olmos Migueláñez, S., & Martínez Abad, F. (2012). Propiedades métricas y estructura dimensional de la adaptación española de una escala de evaluación de competencia informacional autopercibida (IL-HUMASS). *Revista de investigación educativa*, *30*(2), 347-365.

Sánchez-Prieto, J. C., Hernández-García, Á., García-Peñalvo, F. J., Chaparro-Peláez, J., & Olmos-Migueláñez, S. (2019). Break the walls! Second-Order barriers and the acceptance of mLearning by first-year pre-service teachers. *Computers in Human Behavior*, *95*, 158–167. https://doi.org/10.1016/j.chb.2019.01.019

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! Journal of Business Research, 69(10), 3998-4010. https://doi.org/10.1016/j.jbusres.2016.06.007

Schmitt, T. A., Sass, D. A., Chappelle, W., & Thompson, W. (2018). Selecting the "Best" Factor Structure and Moving Measurement Validation Forward: An Illustration. *Journal of Personality Assessment*, *100*(4), 345-362. https://doi.org/10.1080/00223891.2018.1449116

Šupak Smolčić, V. (2013). Salami publication: Definitions and examples. *Biochemia Medica*, *23*(3), 237-241. https://doi.org/10.11613/BM.2013.030

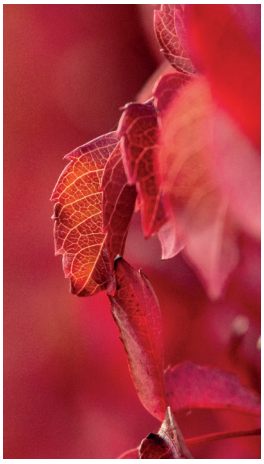
Thomas, H. J., Scott, J. G., Coates, J. M., & Connor, J. P. (2019). Development and validation of the Bullying and Cyberbullying Scale for Adolescents: A multi-dimensional measurement model. *British Journal of Educational Psychology*, *89*(1), 75-94. https://doi.org/10.1111/bjep.12223

Tourón, J., Martín, D., Navarro-Asencio, E., Pradas, S., & Iñigo, V. (2018). Validación de constructo de un instrumento para medir la competencia digital docente de los profesores (CDD). *Revista española de pedagogía*, *76*(269), 25-54. https://doi.org/10.22550/REP76-1-2018-02

Tourón, M., Tourón, J., & Navarro-Asencio, E. (2023). Validez de Constructo de la Escala de Detección de Alumnado con Altas Capacidades para Profesores de Educación Infantil, Gifted Rating Scales (GRS2-P), en una muestra española. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, *29*(2), Article 2. https://doi.org/10.30827/relieve.v29i2.27787

Tourón, M., Tourón, J., & Navarro-Asencio, E. (2024). Validación española de la Escala de Detección de altas capacidades, «Gifted Rating Scales 2 (GRS 2-S) School Form», para profesores. *Estudios sobre Educación*, *46*, 33-55. https://doi.org/10.15581/004.46.002

Información de contacto: Fernando Martínez Abad, Instituto Universitario de Ciencias de la Educación. Universidad de Salamanca.Paseo Canalejas, 169. 37008, Salamanca. E-mail: fma@usal.es



Investigaciones

Un análisis bibliométrico sobre la eficiencia en la educación

A Bibliometric Analysis on Efficiency in Education

https://doi.org/10.4438/1988-592X-RE-2024-406-639

Muammer Maral
https://orcid.org/0000-0002-2055-5711
National Defense University

Resumen

La eficiencia, un componente del rendimiento, se refiere a lo bien que una unidad de decisión utiliza sus recursos en el proceso de producción. La educación es un proceso complejo con múltiples entradas y múltiples salidas. Para aprovechar al máximo los beneficios de la educación, los recursos deben utilizarse de la mejor manera. Esto evitará el despilfarro de recursos y aumentará los resultados. Por estas razones, la eficacia de la educación se ha convertido en los últimos años en uno de los temas más debatidos entre los responsables políticos, los investigadores y los responsables de la toma de decisiones. Este estudio ofrece una revisión exhaustiva de la creciente bibliografía sobre la eficacia educativa. Esta investigación tiene como objetivo revelar la estructura del campo abordando la eficiencia en la literatura educativa desde una perspectiva holística. El estudio realizó un análisis bibliométrico con 1315 publicaciones que abarcan el período comprendido entre 1970 y 2022 obtenidas de las bases de datos SCOPUS y Web of Science. Según los resultados del estudio, la tasa de crecimiento anual de la eficiencia en la literatura educativa es de aproximadamente un 9%. Los últimos 10 años de crecimiento cubren más de la mitad de los 50 años de literatura. Esta investigación reveló las contribuciones al campo en términos de revistas, autores, instituciones y países. Se identificaron las publicaciones más influyentes del campo. Entre 78 países, China y Estados Unidos han realizado la mayor contribución y colaboración. El análisis envolvente de datos, el análisis de frontera estocástica, la educación superior, la productividad y el rendimiento, la eficiencia técnica, la eficiencia de escala, la eficiencia de asignación y la eficiencia de costos son los temas principales del campo. La eficiencia de la investigación,

Recibido: 17/02/2023 Aceptado: 13/10/2023

la clasificación, la calidad, la eficiencia escolar, la competencia, el aprendizaje electrónico y la educación en Europa son los temas de moda en este campo.

Palabras clave: eficiencia, educación, eficiencia en la educación, revisión, análisis bibliométrico.

Abstract

Efficiency, a component of performance, refers to how well a decision-making unit utilises its resources in the production process. Education is a complex process with multiple inputs and multiple outputs. In order to make the most of the benefits of education, resources should be used in the best way. This will prevent waste of resources as well as increase the outputs. For these reasons, efficiency in education has become one of the most frequently discussed issues among policy makers, researchers and decision makers in recent years. This study provides a comprehensive review of the growing literature on educational efficiency. This research aims to reveal the structure of the field by addressing the efficiency in education literature from a holistic perspective. The study conducted a bibliometric analysis with 1315 publications covering the period between 1970 and 2022 obtained from the SCOPUS and Web of Science databases. According to the results of the study, the annual growth rate of efficiency in education literature is approximately 9%. The last 10 years of growth cover more than half of the 50 years of literature. This research revealed the contributions to the field in terms of journals, authors, institutions and countries. The most impactful publications of the field were identified. Among 78 countries, China and the United States have made the largest contribution and collaborations. Data envelopment analysis, stochastic frontier analysis, higher education, productivity and performance, technical efficiency, scale efficiency, allocative efficiency, and cost efficiency are the main themes of the field. Research efficiency, ranking, quality, school efficiency, competition, e-learning, and education in Europe are the trending topics in the field.

Keywords: efficiency, education, efficiency in education, review, bibliometric analysis.

Introducción

La eficiencia es un concepto que expresa lo bien que una unidad de decisión utiliza sus recursos en la producción de bienes y servicios (Kalb, 2010). En el contexto de la educación, la eficiencia, es decir, el

Recibido: 17/02/2023 Aceptado: 13/10/2023

uso eficiente de los recursos, se refiere a la producción del máximo rendimiento educativo utilizando el mínimo de insumos (Johnes et al., 2017). En educación, se puede hablar de eficiencia si los insumos existentes se utilizan de la manera más óptima posible, es decir, si se puede producir el máximo rendimiento. En un sistema ineficiente, existe la posibilidad de aumentar los resultados educativos con un determinado nivel de gasto o de disminuir los recursos educativos utilizados para determinados resultados educativos (Bessent y Bessent, 1980).

La eficiencia es uno de los temas examinados por diferentes disciplinas en la literatura. En la literatura, hay estudios sobre la eficiencia en muchos campos como la economía (Chen et al., 2018; Khodadadipour et al., 2021; Tan, 2006; Vidoli & Ferrara, 2015), la administración (Ang et al., 2021; Holmgren, 2018; Venkadasalam et al., 2020), energía (De Clercq et al., 2019; Moutinho et al., 2020; Quintano et al, 2021), salud (Bayley et al., 2022; Gómez-Gallego et al., 2021; Moreno-Enguix et al., 2018), transporte (Chang et al., 2018; Chao et al., 2018; Park et al., 2018), etc.

Kalirajan y Shand (1999) afirman que la medición de la eficiencia tiene principalmente tres ventajas. En primer lugar, la medición de la eficiencia ofrece la posibilidad de comparar unidades similares. De este modo, se pueden comparar unidades homogéneas entre sí y determinar sus niveles relativos de eficiencia. En segundo lugar, la medición de la eficiencia permite identificar el origen de las diferencias de eficiencia entre unidades. En tercer lugar, la medición de la eficiencia revela algunas implicaciones para mejorar la eficiencia de las unidades.

Teniendo en cuenta estos beneficios de la medición de la eficacia, se puede ver que se han realizado muchos estudios sobre la eficacia en el campo de la educación, al igual que en otras disciplinas. El ámbito de la investigación sobre la eficacia en la educación (EE) es muy diverso. Algunas de las investigaciones se centran en la eficiencia de un determinado nivel educativo en un país (Agasisti & Dal Bianco, 2009; Andersson et al., 2017; Brzezicki, 2020; Dufrechou, 2016). Algunos estudios examinan la eficiencia educativa de más de un país (Agasisti, 2014; Agasisti & Zoido, 2018; Delprato & Antequera, 2021). Además de estos, los investigadores también han realizado estudios que examinan las instituciones educativas públicas (Canal et al., 2015; Kantabutra & Tang, 2010; Turkan & Ozel, 2017; Visbal-Cadavid et al., 2017) y las instituciones educativas privadas (Bayraktar et al., 2013; Shamohammadi & Oh, 2019). Algunos estudios también han examinado la eficiencia de subsistemas

como los departamentos académicos (Anastasiou et al., 2007; Kao & Hung, 2008) y las bibliotecas (Reichmann, 2004; Tavares et al., 2018). Algunos estudios abordan las misiones de enseñanza e investigación de las instituciones de educación superior (Gralka et al., 2019; Jiang et al., 2020; Maral, 2023; Tran et al., 2020).

La literatura muestra que el alcance de la investigación sobre EE es diverso y la literatura tiene una perspectiva rica. Witte y López-Torres (2017) encontraron que la literatura sobre EE está creciendo. Por lo tanto, es necesario examinar estos estudios realizados en diferentes niveles y campos de la educación con una perspectiva holística. Existen algunos estudios en la literatura que son paralelos a este propósito.

Rhaiem (2017) realizó una revisión bibliográfica de la investigación sobre la eficiencia de la investigación académica. Worthington (2001) enumeró los trabajos que utilizaban métodos de estimación de fronteras para medir la eficiencia en la educación. Johnes (2004) describió las técnicas utilizadas para medir la eficiencia y explicó los usos y desventajas de aplicar diferentes métodos en un contexto educativo. Lampe y Hilgers (2015) realizaron un análisis bibliométrico sobre el análisis envolvente de datos y el análisis de frontera estocástica, dos de los métodos de medición de la eficiencia más utilizados. Ferro y D'Elia (2020) revisaron la literatura sobre las variables de input y output utilizadas en la investigación sobre la eficiencia en la educación superior. Witte y López-Torres (2017) proporcionan una descripción detallada de los inputs, outputs y variables no discrecionales utilizadas en la investigación sobre EE. Además, revelan los métodos de medición de la eficiencia utilizados en los estudios junto con sus fuentes. Emrouznejad et al. (2008) examinaron el estado del análisis envolvente de datos a lo largo de 30 años con un punto de vista descriptivo. Villano y Tran (2021) realizaron un análisis de metarregresión para comprender mejor la relación entre el análisis envolvente de datos y la calidad de la investigación sobre eficiencia en la enseñanza superior. Çetin y Maral (2022) examinaron la investigación sobre la eficiencia en la enseñanza superior desde el punto de vista del contenido y la metodología.

Se observa que cada uno de estos estudios ha realizado aportaciones significativas al campo de la eficiencia. Sin embargo, al evaluar estos estudios en su conjunto, se observa que se centran en un área específica en términos de contenido o metodología, o no examinan específicamente todo el campo de la EE. Por lo tanto, es necesario abordar las aplicaciones

Recibido: 17/02/2023 Aceptado: 13/10/2023

de la investigación sobre EE en el ámbito de la educación desde una perspectiva holística con un enfoque macro. Por lo tanto, este estudio tiene como objetivo examinar la investigación sobre la eficiencia en la educación desde una perspectiva holística, explorar la estructura intelectual del campo, revelar patrones en la literatura y contribuir al avance del campo de formas nuevas y significativas. Para lograr este objetivo, el estudio trata de responder a las siguientes preguntas:

1. ¿Cuál es la bibliografía básica de la investigación sobre EE? Para revelar la estructura intelectual de la investigación resultante, los investigadores se centraron en los siguientes datos bibliométricos clave: (1) Número de publicaciones y citas a lo largo del tiempo, (2) Publicaciones más influyentes, (3) Revistas más influyentes y productivas, (4) Autores más influyentes y productivos, (5) Instituciones con mayor contribución, (6) Países con mayor contribución y (7) Palabras clave más utilizadas.
2. ¿Qué temas han surgido en la bibliografía sobre EE? Se realizó un "análisis de mapa temático" para revelar los temas principales de los estudios.
3. ¿Cuáles son los temas de tendencia en la literatura sobre EE? Se llevó a cabo un análisis de los "temas de tendencia" para identificar los temas de tendencia que han surgido en la literatura sobre EE en los últimos años.
4. ¿Cuál es la interacción social entre países en la literatura sobre EE? Se llevó a cabo un "análisis de los países asociados" para revelar la interacción social y las conexiones entre países y sus efectos en el desarrollo del campo de la literatura sobre EE.

Método

Este estudio utilizó un método bibliométrico para revisar la literatura de EE de forma exhaustiva y acumulativa, determinar el aspecto general y las tendencias de las publicaciones en la literatura de EE, revelar las interacciones sociales y desvelar el mapa científico de la literatura. Los estudios bibliométricos intentan sintetizar patrones de producción de conocimiento a través del análisis de datos bibliográficos asociados a un cuerpo de documentos relevantes (Van Eck & Waltman, 2014; Zupic

& Čater, 2015). Este tipo de revisión sistemática revela las tendencias evolutivas de la literatura de un campo a lo largo del tiempo y proporciona fundamentos empíricos para el camino a seguir (Hallinger, 2021). El volumen de la investigación científica ha aumentado considerablemente en los últimos años. Esto ha hecho que a los investigadores les resulte cada vez más difícil seguir la literatura relevante en su campo. Este hecho hace necesario el uso de métodos bibliométricos cuantitativos que puedan manejar esta vasta riqueza de datos, estimar el impacto de estudios importantes y explorar la estructura subyacente de la literatura en un campo (Zupic & Čater, 2015).

Identificación de recursos

Este estudio utilizó las bases de datos Web of Science Core Collection (WOS) y Scopus, que cubren una parte significativa de la literatura mundial y son las más utilizadas en los estudios bibliométricos. Los estudios que recurren al análisis bibliométrico suelen utilizar sólo una de estas dos bases de datos. Los investigadores decidieron utilizar estas dos bases de datos juntas para cubrir la bibliografía sobre EE de forma más exhaustiva.

Los criterios de inclusión de las fuentes son los siguientes: (1) El tema de investigación es la eficiencia, (2) La investigación es en el campo de la educación, (3) La investigación cae en una de las categorías de un artículo, artículo, revisión de la literatura, capítulo de libro, y el libro, (4) La investigación fue publicada antes del 12 de septiembre de 2022. Los criterios de exclusión son los siguientes (1) Investigación sobre la eficiencia en cualquier campo que no sea la educación, (2) Estudios que no sean artículos, ponencias, revisiones bibliográficas, libros y capítulos de libros. La razón de incluir en este estudio artículos, capítulos de libros y libros, lo que se denomina literatura gris, es evitar que se excluyan del análisis las fuentes principales e influyentes del tema. Además, en este estudio no se ha utilizado ningún filtro temporal. Los recursos obtenidos abarcan el periodo comprendido entre 1970 y 2022.

La selección de las palabras clave adecuadas es muy importante en los estudios bibliométricos y esto afecta directamente a los resultados (Sweileh, 2018). La estrategia de búsqueda final elegida tras muchos intentos de búsqueda de diferentes combinaciones de palabras clave se

Recibido: 17/02/2023 Aceptado: 13/10/2023

presenta en **el Apéndice.** El investigador desarrolló una estrategia de búsqueda exhaustiva para garantizar que ninguna de las fuentes relevantes se pasara por alto en la formulación de la estrategia de búsqueda. Cuanto más específicos y limitados sean los conceptos utilizados en las bases de datos, menor será el número de publicaciones que aparezcan. Sin embargo, si el tema de la eficacia no tiene esta cualidad y aun así una persona acota la estrategia de búsqueda, puede pasar por alto muchos recursos relevantes. También existe el riesgo de perderse recursos importantes sobre el mismo tema de revistas fundamentales externas. Por lo tanto, este estudio desarrolló una estrategia de búsqueda exhaustiva. La revisión preliminar de la bibliografía mostró que la investigación sobre EE se llevaba a cabo en muchas disciplinas. Para acotar los resultados de la búsqueda se utilizó el filtrado por materias de las bases de datos. Sin embargo, los investigadores filtraron los campos para excluir las áreas temáticas más alejadas de la educación (por ejemplo, farmacia, química o neurociencia). Sin embargo, aunque esto es muy poco probable, puede haber dado lugar a que se pasaran por alto algunas investigaciones sobre la eficiencia en la educación en estas áreas.

En la estrategia de búsqueda, TI corresponde a las búsquedas en el título, AK a las búsquedas en las palabras clave del autor y AB a las búsquedas en el resumen. El investigador buscó la frase "efficienc*" en el título y en las palabras clave del autor. La presencia de este término en el título y en las palabras clave del autor facilitó la recuperación de investigaciones relacionadas con la eficiencia. Sin embargo, este concepto no se buscó en el resumen. La razón principal es que, si la bibliografía es extensa, al buscar en el resumen pueden incluirse en el conjunto de datos muchos falsos positivos, es decir, estudios irrelevantes. Aunque todas las publicaciones podrían haberse analizado individualmente con los criterios de inclusión y exclusión, esto se evitó debido al elevado número de publicaciones que debían analizarse. Esto puede haber dado lugar a que algunos estudios relevantes no se incluyeran en el conjunto de datos. Dado que en la bibliografía sobre el tema de la eficiencia hay investigaciones de muchas disciplinas, también se incluyeron en el título conceptos relacionados con la educación (por ejemplo, educación, universidad, enseñanza superior, etc.) para distinguir en la medida de lo posible los relacionados con la educación. Sin embargo, aunque esto nos acerca al campo de la EE, también incluyó en los datos muchos estudios distintos del concepto de eficiencia económica. Para evitarlo,

los investigadores buscaron conceptos específicos de eficiencia (por ejemplo, análisis envolvente de datos, análisis de frontera estocástica, eficiencia técnica, etc.) en el título, el resumen y las palabras clave del autor. Esta estrategia de búsqueda permitió al investigador acercarse a los estudios de interés. Los investigadores decidieron utilizar la estrategia de búsqueda de este modo porque si se restringía más la estrategia de búsqueda podrían quedar excluidos estudios pertinentes.

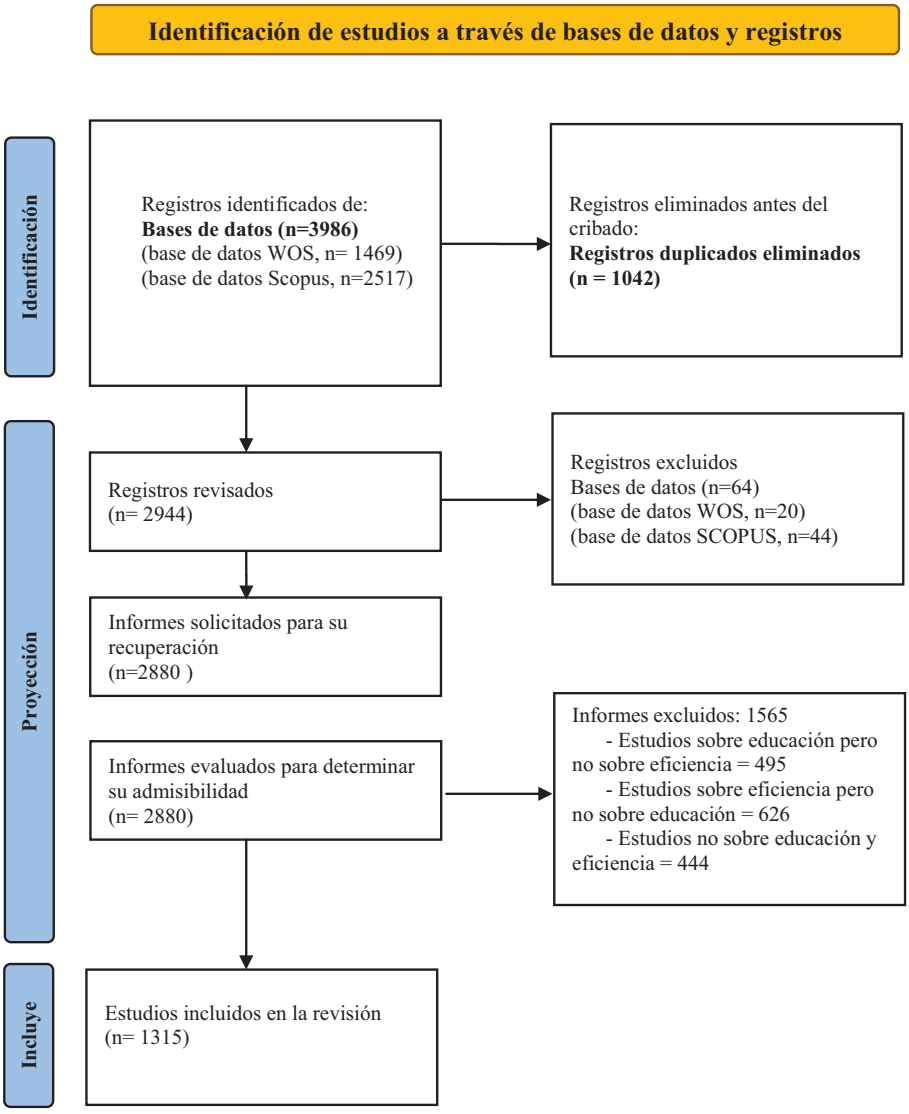
Extracción y análisis de datos

Esta revisión siguió las directrices PRISMA utilizadas habitualmente para informar sobre revisiones sistemáticas de la investigación (Page et al., 2021). La Figura I presenta el diagrama de flujo PRISMA seguido en este estudio. Con la estrategia de búsqueda determinada, se realizaron búsquedas separadas en WOS y SCOPUS (el 12 de septiembre de 2022). La base de datos WOS proporcionó 1469 documentos y la base de datos SCOPUS proporcionó 2517 documentos. Dado que 1042 del total de 3986 documentos obtenidos eran duplicados, los investigadores los eliminaron del conjunto de datos. De los 2944 documentos restantes, 64 se excluyeron del conjunto de datos por no cumplir el tipo de documento especificado en los criterios de inclusión. De los 2.880 documentos restantes, 1.565 se excluyeron del conjunto de datos por ser "falsos positivos" irrelevantes. En la fase final, quedaban 1315 documentos por incluir en el análisis.

Los datos bibliométricos asociados a los 3986 documentos se exportaron de WOS a un archivo de texto sin formato. Los datos de SCOPUS se descargaron en dos partes, se fusionaron y se guardaron como archivo ".bib". Los documentos obtenidos de WOS y SCOPUS se convirtieron en un único documento de datos bibliométricos de MS Excel utilizando el programa R Studio. Estos documentos de datos contienen información descriptiva como autor, título, datos de publicación, datos de citación, resumen, etc. El autor revisó el documento de datos MS Excel generado para eliminar las publicaciones irrelevantes. Fue necesario eliminar los estudios no relevantes para el tema una vez identificados los estudios. Por ejemplo, una publicación puede contener palabras clave utilizadas en la búsqueda. Sin embargo, este estudio puede no ser relevante para el tema de interés. En tales casos, estos estudios pueden generar "falsos positivos" (Linnenluecke et al., 2020). Para eliminar

Recibido: 17/02/2023 Aceptado: 13/10/2023

FIGURA I. Diagrama de flujo PRISMA



Fuente: Elaboración propia.

los falsos positivos, se identificaron todas las publicaciones relevantes e irrelevantes para EE. El autor revisó este documento de datos y lo envió a tres expertos independientes en la materia para su revisión.

Recibido: 17/02/2023 Aceptado: 13/10/2023

Se introdujeron cambios en el documento de datos en función de las opiniones de los expertos de campo. En los casos en que las opiniones de los expertos eran contradictorias, los investigadores se reunieron con los expertos y acordaron los documentos que debían incluirse y excluirse. Por último, se obtuvo un amplio corpus de datos de muchas disciplinas relacionadas con el tema estudiado.

En el análisis bibliométrico, uno de los puntos importantes para la exactitud y fiabilidad del análisis es la limpieza de los datos. Problemas como valores duplicados, errores ortográficos y tipográficos y falta de información importante (año, nombre del autor, nombre de la revista, etc.) deben corregirse en los documentos recuperados de las bases de datos. Porque, aunque la mayoría de los datos bibliométricos son fiables, a veces pueden contener varias versiones del mismo estudio. Puede haber diferencias en la ortografía de autores y revistas (Zupic & Čater, 2015). En este estudio, la limpieza de datos fue un proceso meticuloso. En primer lugar, se eliminaron los trabajos duplicados. En segundo lugar, se completó la información importante que faltaba accediendo a bases de datos o, en caso necesario, al texto completo del estudio. En tercer lugar, se examinaron las diferencias ortográficas (por ejemplo, palabras en singular y plural como efficiency/efficiencies, grafías con guión como higher education/higher-education, o abreviaturas como DEA/data envelopment analysis, etc.). En cuarto lugar, antes del análisis temático utilizado para revelar la estructura temática del campo de la EE, se eliminaron las palabras clave de autor. En esta fase, se combinaron diferentes palabras con el mismo significado mediante la preparación de un archivo de texto de "texto sinónimo" utilizado en el software R Bibliometrix. De este modo, se agruparon palabras dispersas con el mismo significado. Otro aspecto importante del análisis temático es la eliminación de palabras innecesarias y sin sentido del análisis para revelar los temas del campo estudiado. Por ejemplo, expresiones muy generales como "education, input, and output" o palabras que no tienen sentido para el análisis temático como "i21, 2009" se excluyeron del análisis preparando un archivo de texto "remove text" utilizado en el software R Bibliometrix.

Para el análisis de los datos se utilizó el programa MS Excel, los paquetes de software Bibliometrix y Biblioshiny en la herramienta R. El paquete R Bibliometrix es una herramienta de software de código abierto que proporciona un conjunto de herramientas para la investigación cuantitativa en metodología científica. El software R incluye un gran

Recibido: 17/02/2023 Aceptado: 13/10/2023

número de algoritmos estadísticos eficientes. Las operaciones numéricas de alta calidad y la disponibilidad de herramientas de visualización de datos se encuentran entre las características fuertes de los lenguajes R en comparación con otros lenguajes de la informática científica (Aria & Cuccurullo, 2017). Biblioshiny es un desarrollo secundario del paquete Shiny basado en Bibliometrix en el lenguaje R. Biblioshiny incluye el código central de Bibliometrix y ofrece análisis de datos basados en la web. Biblioshiny permite realizar mediciones científicas y análisis visuales (Xie et al., 2020).

Este estudio pretende responder fundamentalmente a cuatro preguntas. El investigador realizó análisis bibliométricos básicos para responder a la primera pregunta del estudio. Como medida, el investigador analizó los autores, documentos, revistas, instituciones y países según el número de publicaciones, el total de citas y las citas por publicación. La relación entre el número de documentos y el número total de citas es una buena medida del impacto de las publicaciones y los colaboradores (Saravanan et al., 2022).

El investigador realizó un análisis temático para responder a la segunda pregunta. El análisis temático es un método utilizado para proporcionar visualización e interpretación del tema e identificar las tendencias actuales en un periodo de tiempo determinado (Cobo et al., 2011). Este enfoque crea representaciones gráficas que resumen automáticamente los temas principales de una publicación. También se pueden caracterizar los temas emergentes según su estructura y papel en la red. El análisis temático revela un diagrama estratégico. Este diagrama tiene dos dimensiones, intensidad y centralidad, y permite destacar cuatro tipos diferentes de temas en función del cuadrante en el que se encuentren. Los valores altos de centralidad e intensidad identifican temas populares relevantes para estructurar el marco conceptual del campo. Los valores de centralidad alta e intensidad baja definen los temas clave importantes para el área. Los valores de baja centralidad y baja intensidad identifican temas aún no desarrollados o marginalmente interesantes para el área. Los valores de baja centralidad y alta densidad identifican temas muy desarrollados, pero de nicho para el área, con potencial de desarrollo (Aria et al., 2022).

Para responder a la tercera pregunta se utilizó el análisis de tendencias. El análisis de tendencias es un análisis que revela los últimos estudios de tendencias en un campo. Esto determina la dirección del campo y orienta a los autores para futuras investigaciones.

Para responder a la cuarta pregunta, se realizó un análisis de coautoría por países. Este análisis pretende revelar la estructura social del área y sus interacciones. Dado que un artículo es publicado por más de un autor, institución o país, se forma una relación entre ellos (Lu & Wolfram, 2012). La coautoría de publicaciones científicas examina las redes sociales que forman científicos, instituciones y países al colaborar en publicaciones científicas (Adamides & Karacapilidis, 2006). El análisis de las redes de colaboración permite conocer los posibles socios de investigación de un campo para la publicación de futuras investigaciones (Xu et al., 2022). La coautoría se considera una medida de la colaboración. Esta medida refleja los vínculos sociales con mayor intensidad que otras medidas de colaboración. Además, los datos bibliométricos también contienen información sobre las instituciones y los países de los autores, por lo que también se pueden examinar las redes de colaboración entre instituciones y países (Zupic & Čater, 2015). Este estudio realizó un análisis de coautoría por países para revelar la red de interacción social del campo.

Resultados

El resumen de las publicaciones sobre EE

La Tabla I presenta la visión general de las publicaciones sobre EE. Las publicaciones sobre EE abarcan el periodo comprendido entre 1970 y 2022 y hay un total de 1315 documentos. La tasa de crecimiento anual de las publicaciones es del 8,92%. La bibliografía incluye contribuciones de 78 países y un total de 2349 autores.

La Figura II presenta el número de publicaciones sobre EE y el número medio de citas por año. Se observa que el número de publicaciones sobre EE ha aumentado considerablemente a lo largo de los años. No hubo un número elevado de publicaciones en el periodo comprendido entre los años setenta y 2002. Especialmente desde 2002, se ha producido un periodo de aumento del número de publicaciones. En 2015, se produjo un descenso en el número de publicaciones, pero el aumento continuó en los años siguientes. Esta tendencia ha alcanzado su punto máximo en los últimos 10 años, con un 68% del total de publicaciones. Por otra parte, el número medio de citas por año no es regular y fluctúa constantemente al alza y a la baja.

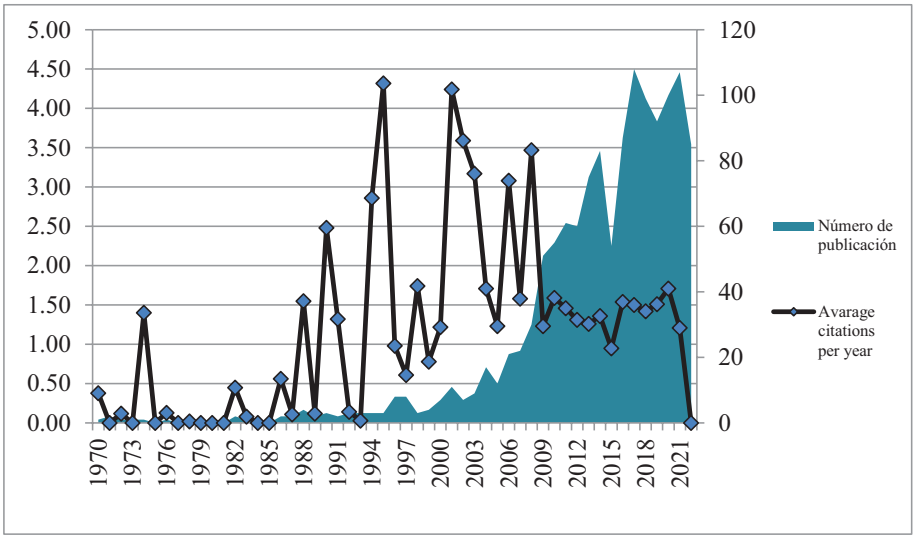
Recibido: 17/02/2023 Aceptado: 13/10/2023

TABLA I. Información básica sobre las publicaciones

Duración	1970-2022
Publicaciones	1315
Tasa de crecimiento anual	8,92
Publicación Edad media	8,07
Promedio de citas por publicación	12,7
Referencias	32021
Palabras clave del autor	2368
Palabra clave plus	1824
Autores	2349
Países	78
Documentos de autor único	267
Coautores por documento	2,5
Coautorías internacionales %.	10,65

Fuente: Elaboración propia.

FIGURA II. Publicación de literatura de EE y medias de citas



Fuente: Elaboración propia.

Recibido: 17/02/2023 Aceptado: 13/10/2023

Las publicaciones más influyentes

La Tabla II presenta la lista de las diez publicaciones más citadas en la literatura sobre EE. La publicación con más citas totales (CT) y citas totales por año (CTY) fue la investigación de Emrouznejad et al. (2008). Este estudio examinó un historial bibliográfico de 30 años de la literatura del análisis envolvente de datos. Uno de los resultados importantes de la investigación es que la bibliografía sobre el uso del análisis envolvente de datos ha aumentado significativamente durante este periodo. Avkiran (2001) y Abbott y Doucouliagos (2003), que se encuentran entre los tres primeros según el número total de citas, examinaron la eficiencia técnica y de escala de las universidades en Australia.

TABLA II. Publicaciones más influyentes en la literatura sobre EE

Autor(es)	**Publicación**	**TC**	**TCY**
Emrouznejad, A., Parker, B. R., & Tavares, G. (2008)	Evaluación de la investigación en eficiencia y productividad: una encuesta y análisis de los primeros 30 años de literatura académica en la DEA.	703	46,87
Avkiran, N. K. (2001)	Investigación de la eficiencia técnica y de escala de las universidades australianas mediante el análisis envolvente de datos	359	16,32
Abbott, M., & Doucouliagos, C. (2003)	La eficiencia de las universidades australianas: Un análisis envolvente de datos	349	17,45
Johnes, J. (2006).	El análisis envolvente de datos y su aplicación a la medición de la eficiencia en la enseñanza superior	309	18,18
Thursby, J. G., & Kemp, S. (2002)	Crecimiento y eficiencia productiva de las licencias universitarias de propiedad intelectual	284	13,52
Auranen, O., y Nieminen, M. (2010).	Financiación de la investigación universitaria y resultados de las publicaciones: comparación internacional	260	20,00
Beasley, J. E. (1995).	Determinar la eficacia de la enseñanza y la investigación	255	9,11
Beasley, J. E. (1990)	Comparación de departamentos universitarios	216	6,55
Anderson, T. R., Daim, T. U., & Lavoie, F. F. (2007)	Medir la eficacia de la transferencia de tecnología universitaria	186	11,63
Afonso, A., & Aubyn, M. S. (2006).	Eficiencia de la enseñanza secundaria en distintos países: un análisis semiparamétrico con insumos no discrecionales	160	9,41

TC: total de citas, TCY: total de citas por año. Las publicaciones están ordenadas por TC. Fuente: Elaboración propia.

Las revistas más productivas e influyentes en la literatura de EE.

En la Tabla III figuran las diez revistas más prolíficas e influyentes de la literatura sobre EE. La revista con mayor número de publicaciones es "Revista de Economía de la Educación". Esta revista publica investigaciones sobre políticas y finanzas educativas, producción de capital humano y rendimientos del capital humano. El hecho de que la revista publique temas relacionados con la política y las finanzas de la educación es un factor importante para que sea la revista con mayor número de publicaciones en el campo de la EE. Esta revista también ocupa el primer lugar en cuanto al número total de citas. La segunda revista que contribuye a este campo es "Ciencias de la Planificación Socioeconómica". Ocupa el segundo lugar en cuanto a número de publicaciones y citas totales. La revista que destaca en la lista es "Revista de la Sociedad de Investigación Operativa". Esta revista, que publica estudios sobre la toma de decisiones y la investigación operativa, tiene el valor CP más alto. Esto demuestra que los estudios publicados tienen un alto impacto y esto puede aumentar la importancia de la revista en el campo de la EE en el futuro.

TABLA III. Las revistas más productivas e influyentes

Rango	Revista	NP	TC	CP	h indice
1	Economics of Education Review	34	1684	49,5	21
2	Socio-Economic Planning Sciences	33	1438	43,6	13
3	Applied Economics	27	566	21,0	13
4	Education Economics	22	817	37,1	16
5	European Journal of Operational Research	19	868	45,7	13
6	Journal of The Operational Research Society	17	866	50,9	13
7	Scientometrics	14	347	24,8	9
8	İnternational Journal of Educational Development	12	69	5,8	4
9	Journal of Productivity Analysis	12	194	16,2	9
10	Higher Education	10	336	33,6	8

NP: número de publicaciones, CT: total de citas, CP: media de citas por publicación. Las revistas se enumeran según el número de publicaciones. Fuente: Elaboración propia.

Recibido: 17/02/2023 Aceptado: 13/10/2023

TABLA IV. Los autores más productivos e influyentes

Rango	Autores	NP	TC	CP	h_indice
1	Agasisti, T.	38	1071	28,2	20
2	Johnes, J.	18	1047	58,2	13
3	Johnes, G.	16	618	38,6	11
4	De W. K.	15	589	39,3	12
5	Prior, D.	14	177	12,6	7
6	Tran, C.	11	85	7,7	5
7	Ruggiero, J.	10	237	23,7	8
8	Villano, R.	10	86	8,6	5
9	Barra, C.	9	134	14,9	7
10	Zotti, R.	9	134	14,9	7

NP: número de publicaciones, CT: total de citas, CP: media de citas por publicación. Los autores se enumeran según el número de publicaciones. Fuente: Elaboración propia.

Los autores más prolíficos e influyentes de la literatura de EE

La Tabla IV presenta la lista de autores que han hecho contribuciones significativas a la literatura sobre EE. Los diez primeros autores figuran en la lista según el número de publicaciones. Agasisti, T. encabeza la lista con 38 publicaciones. En cuanto al número de publicaciones, se observa que este autor ha publicado un número considerablemente mayor de publicaciones en comparación con otros autores. En segundo lugar, se encuentra Johnes, J. La característica llamativa de este autor es que sus publicaciones son muy citadas. Tienen el valor CP más alto entre todos los demás autores. Otros autores entre los cuatro primeros tanto en términos de NP como de CP son Johnes, G. y De, W.K.

Instituciones que más contribuyen a la literatura sobre EE

La Tabla V presenta la lista de instituciones que han hecho contribuciones significativas a la literatura sobre EE. Se observa que la Universidad de Lancaster es la que más ha contribuido a la literatura sobre EE. Le siguen la Universidad Politécnica de Milán y la Universidad de Maastricht. Entre

Recibido: 17/02/2023 Aceptado: 13/10/2023

TABLA V. Instituciones que más contribuyen

Rango	Afiliación	NP	País
1	Universidad de Lancaster	23	Reino Unido
2	Universidad Politécnica de Milán	17	Italia
3	Universidad de Maastricht	15	Países Bajos
4	Universidad Autónoma de Barcelona	13	España
5	Universidad de Granada	13	España
6	Universidad de Nueva Inglaterra	13	Australia
7	Universidad Complutense de Madrid	12	España
8	Universidad de Economía y Empresa de Praga	12	Chequia
9	Universidad Católica de Lovaina	11	Bélgica
10	Universidad de Extremadura	9	España

Fuente: Elaboración propia.

los principales contribuyentes, hay cuatro instituciones de España. Entre las instituciones de los diez primeros puestos, todos los países, excepto Australia, se encuentran en Europa.

Países con mayor contribución a la literatura sobre EE.

La Figura III presenta la distribución del número de publicaciones de 78 países que han contribuido a la literatura sobre EE. El país que más ha contribuido a la literatura sobre EE es China, con 323 publicaciones. EE.UU. ocupa el segundo lugar con 229 publicaciones. En tercer lugar, se encuentra España con 164 publicaciones, en cuarto lugar, Brasil con 117 publicaciones y en quinto lugar Italia con 112 publicaciones. A estos países les siguen el Reino Unido, Australia, Alemania, Chequia y Rusia. Se observa que los países que menos contribuyen son los del continente africano.

FIGURA III. Los países que más contribuyen

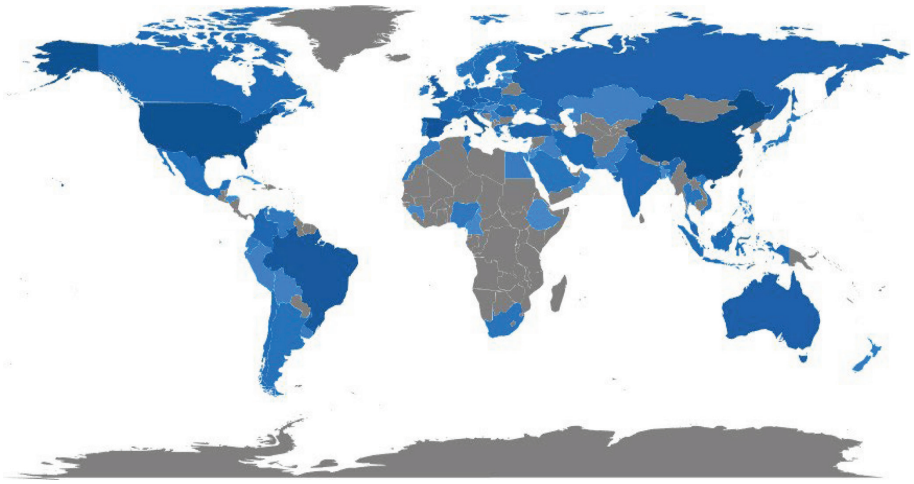


FIGURA IV. Conceptos básicos



El tamaño de letra de las palabras es proporcional a su frecuencia de aparición en las palabras clave). Fuente: Elaboración propia con R Bibliometrix..

Recibido: 17/02/2023 Aceptado: 13/10/2023

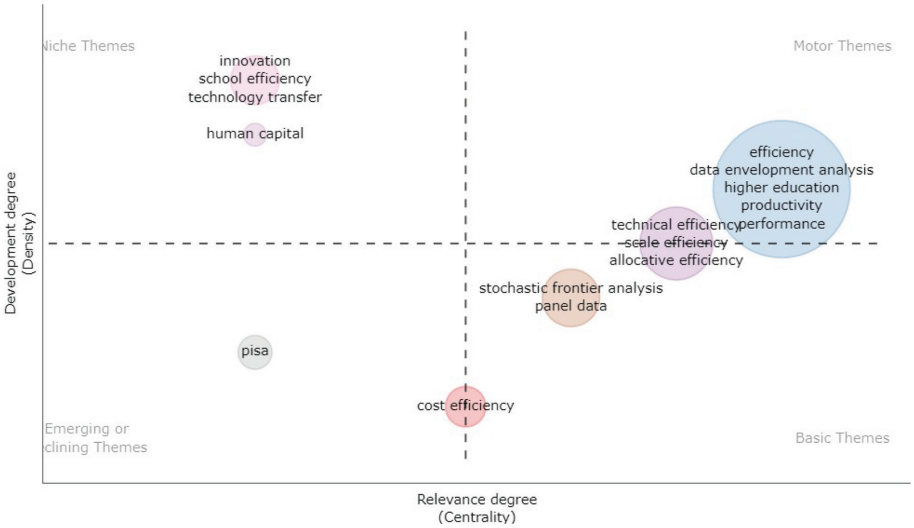
Los conceptos básicos de la literatura sobre EE

Se realizó un análisis de "Nube de palabras" para revelar la estructura de los conceptos en la bibliografía sobre EE. Este análisis se basó en las palabras clave de los autores. La Figura IV presenta los resultados del análisis. Las palabras que más se repiten en la literatura sobre EE son "eficiencia", "análisis envolvente de datos (DEA)", "enseñanza superior", "economía de la educación" y "rendimiento". El DEA es uno de los métodos no paramétricos utilizados para medir la eficiencia de las unidades de decisión. Se puede considerar el hecho de que DEA sea el concepto más utilizado como un indicador de que es uno de los métodos más utilizados en la medición de la eficiencia. Aunque en la literatura sobre EE se han realizado estudios en diferentes niveles educativos, se observa que hay más estudios relacionados con la medición de la eficiencia en la "educación superior". La presencia del término "eficiencia de la investigación" en el mapa es uno de los indicadores de ello.

La visión de los temas en la literatura sobre EE

Para responder a la segunda pregunta de la investigación se realizó un análisis temático. La Figura V presenta los resultados del análisis. En la literatura sobre EE, los estudios sobre educación superior destacan entre los temas populares (temas motores) relacionados con la estructuración del marco conceptual del campo (Abbott & Doucouliagos, 2003). En este tema, cabe destacar que la DEA como método es una herramienta importante para medir la eficiencia (Abbott & Doucouliagos, 2003; Avkiran, 2001; Canal et al., 2015). Además, este tema está bordeado por los "temas centrales" que revelan los conceptos que sustentan el campo y se relacionan con las edades básicas de construcción de la literatura de EE. Otro tema dentro de los temas motores son los tipos de eficiencia. Este tema incluye la eficiencia técnica, la eficiencia de escala, la eficiencia distributiva y la eficiencia de costes. Estos cuatro tipos de actividades se encuentran entre los principales tipos de actividades frecuentemente vistos en la literatura (Andersson et al., 2017; Avkiran, 2001; Turkan & Ozel, 2017). El "análisis de frontera estocástica" es un método paramétrico de medición de la eficiencia que se utiliza con frecuencia en la medición de la eficiencia, aunque no tanto como el DEA (Khodadadipour et al.,

FIGURA V. Visión temática de la literatura sobre EE



Nota: Frecuencia mínima de clusters:5, número de etiquetas: 5, número de palabras: 400. Fuente: Elaboración propia con R Bibliometrixa.

2021). Este método se encuentra dentro del área central de la literatura sobre EE. Existen muchos estudios sobre la medición de la eficiencia con datos de panel (Agasisti & Dal Bianco, 2009; Canal et al., 2015). Este tipo de datos incluye datos que abarcan más de un periodo de tiempo y la medición de la eficiencia se produce con estos datos. De este modo, se puede analizar el cambio en la eficiencia de las unidades de toma de decisiones en diferentes periodos de tiempo.

El área en la que aparece el concepto "PISA" en el mapa temático hace referencia a un tema que ha surgido recientemente o que está perdiendo gradualmente su influencia. Recientemente, en la literatura de EE han surgido estudios en los que se realizan comparaciones entre países de la eficiencia considerando el rendimiento de los países en los exámenes PISA (Agasisti & Zoido, 2018; Delprato & Antequera, 2021). Por lo tanto, las comparaciones de eficiencia entre países basadas en PISA tienen el potencial de convertirse en objeto de más investigaciones con el tiempo.

Por último, los temas emergentes y de nicho son "innovación", "eficiencia escolar", "transferencia de tecnología" y "capital humano".

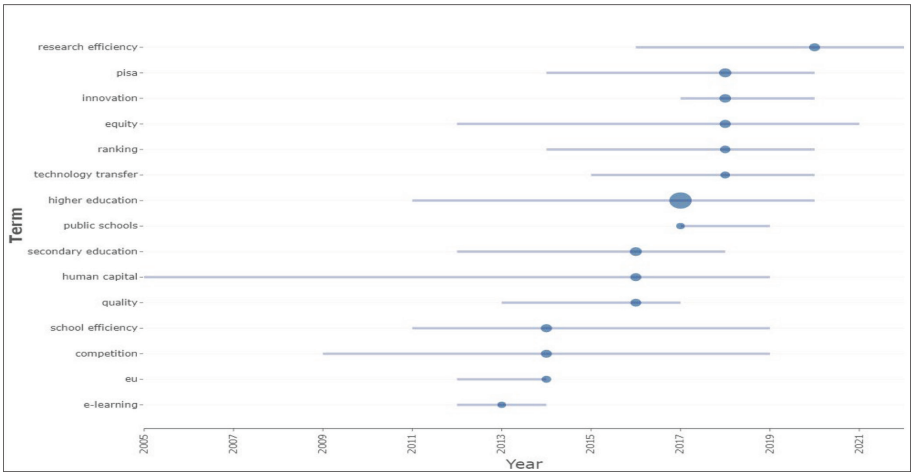
Recibido: 17/02/2023 Aceptado: 13/10/2023

Este tema analiza las actividades basadas en el rendimiento académico de las escuelas primarias y secundarias de una región específica (Agasisti y Zoido, 2018; Barra y Zotti, 2017; Delprato y Antequera, 2021). Los estudios realizados bajo este tema también incluyeron comparaciones de la eficiencia de las escuelas en regiones con diferentes niveles de desarrollo. Dentro de este tema, algunos estudios que analizan la relación entre las universidades y la innovación desde diferentes perspectivas (Shi, et al., 2020). La eficiencia tecnológica, científica y de innovación regional de las universidades, la contribución de las universidades a la innovación regional y la relación entre ellas, y el efecto de la cooperación industria-universidad en la eficiencia de la innovación son algunos de los temas examinados (Agasisti et al., 2019a; Barra & Zotti, 2018; Gong et al., 2022). Un campo importante de la eficiencia de las universidades es la transferencia y aplicación de los conocimientos adquiridos en la universidad a la industria. Por lo tanto, la eficiencia de la transferencia de tecnología también se analiza en el contexto de diferentes países (Agasisti et al., 2019a; Anderson et al., 2007). La relación entre la eficiencia de la transferencia de tecnología de las universidades y el espíritu empresarial también se encuentra entre las cuestiones examinadas. Además de estos, los estudios sobre el capital humano en las universidades también son uno de los temas que se han investigado en los últimos años. La evaluación de la eficiencia de producción de los recursos humanos de las universidades, la evaluación de la eficiencia del personal, la gestión de los recursos humanos y el desarrollo de la eficiencia de este recurso pasan a primer plano. La comparación de la eficiencia de los distintos métodos de aprendizaje en el desarrollo de los recursos humanos y la relación entre la eficiencia de la educación y el nivel de desarrollo humano son algunos de los temas de nicho que destacan en la bibliografía.

Temas de actualidad en la literatura sobre EE

El estudio llevó a cabo un análisis de "Temas de tendencia" para identificar los temas de investigación de tendencia en la literatura sobre EE. La Figura VI presenta los resultados del análisis. Cuando se examina la figura, se puede ver que hay diferentes temas que han sido tendencia en la literatura sobre EE en los últimos años. En los últimos años, la atención se ha centrado en los estudios que miden la eficacia

FIGURA VI. Temas de actualidad en la literatura sobre EE



El tamaño del círculo indica la proporción con respecto al total de publicaciones, y la longitud de la línea indica el alcance del tema examinado. Fuente: Elaboración propia con R Bibliometrix.

investigadora de las universidades. La determinación de la eficiencia educativa de los países basada en los exámenes PISA (Agasisti & Zoido, 2018; Delprato & Antequera, 2021), la eficiencia de innovación de las universidades (Shi, et al., 2020), la eficiencia de las escuelas (Agasisti & Zoido, 2018), la eficiencia de transferencia de tecnología de las universidades (Agasisti et al., 2019a; Anderson et al., 2007), y el capital humano (Barra & Zotti, 2017) se encuentran entre los temas de nicho en el análisis temático anterior. Por lo tanto, estos nichos temáticos también se han convertido en trending topics en los últimos años. Entre los trending topics, el tema más estudiado y dominante en la literatura es el estudio de la eficiencia de las instituciones de educación superior. Además, la eficiencia de la educación financiada con fondos públicos, como la eficacia de las instituciones de educación superior de propiedad estatal, las universidades, las universidades nacionales, las bibliotecas pertenecientes a universidades públicas y la eficacia de los gastos estatales, también son trending topics.

Uno de los temas de moda en la literatura sobre EE es la calidad. La evaluación eficiente de la calidad del aprendizaje de los estudiantes, la calidad de la educación empresarial, la calidad de la enseñanza y la

Recibido: 17/02/2023 Aceptado: 13/10/2023

investigación, y la investigación eficiente sobre el impacto de la calidad del profesorado en el rendimiento de los estudiantes son algunos de los temas examinados (Arbano et al., 2022; Lee & Johnes, 2022). Además, la investigación sobre la calidad de los programas educativos en la enseñanza superior y la investigación de la eficiencia basada en variables de calidad se encuentran entre los temas de tendencia en la literatura de la EE.

En la literatura sobre EE, algunos estudios examinan la competencia de escuelas y universidades en función de la eficiencia (Liu y Xu, 2017). Entre los temas examinados se encuentran el impacto de la competencia entre universidades en la eficiencia y el impacto de la competencia entre escuelas en la eficiencia. Además, la relación entre la competencia y la eficiencia de las instituciones regionales de educación superior, el impacto de la eficiencia financiera de las universidades en la competitividad nacional y el impacto de la cooperación escolar en la eficiencia han sido objeto de estudio recientemente.

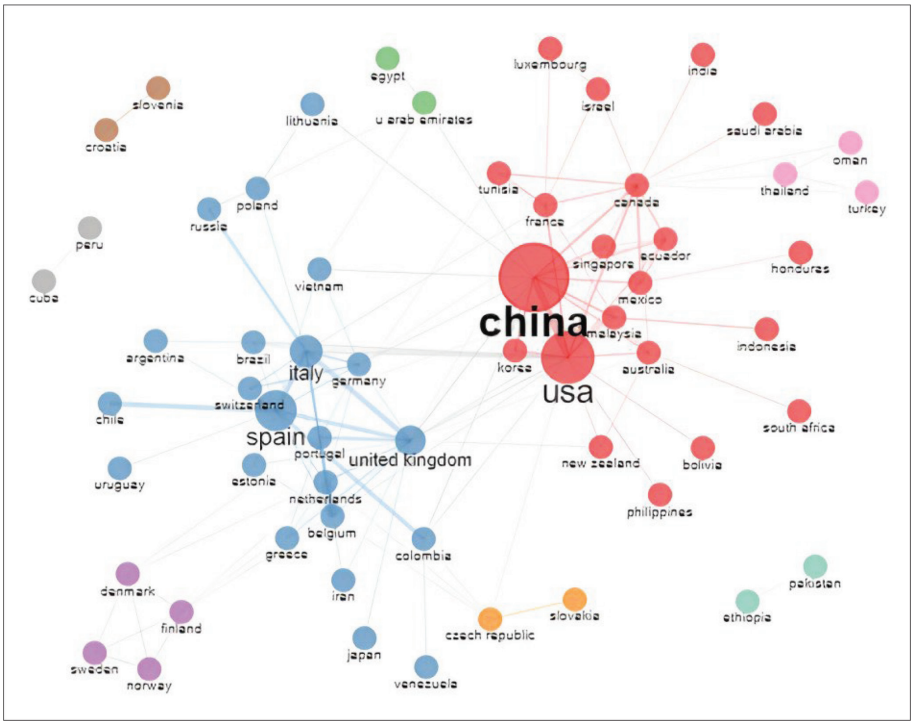
En la literatura de EE, los países europeos y la eficiencia del sistema educativo europeo son uno de los principales intereses de los autores (Agasisti et al., 2019b). El examen de la eficiencia del sistema europeo de educación superior y el análisis comparativo de la eficiencia educativa de diferentes países de Europa se encuentran entre estos temas. Áreas específicas de investigación incluyen la eficiencia de la investigación y el desarrollo del sistema educativo europeo o la eficiencia de los fondos estructurales europeos.

Uno de los temas de moda en la literatura sobre EE son las aplicaciones del e-learning (Chang, 2013). En la bibliografía se examinan, entre otros temas, la eficiencia de las universidades de educación a distancia, la eficiencia de los sistemas de aprendizaje electrónico y la comparación de la eficiencia de los sistemas tradicionales y de aprendizaje electrónico. Además de estos, también hay áreas de estudio como el examen del impacto del uso de aplicaciones móviles digitales de e-learning en la eficiencia de la escuela.

Red de colaboración entre países en la literatura sobre EE

Para revelar la red de colaboración de los países que contribuyen a la literatura sobre EE, se realizó un análisis de la red de colaboración de

FIGURA VII. Red de colaboración entre países



El tamaño del círculo indica la extensión de la colaboración y el grosor de las líneas entre países indica la intensidad de la colaboración. El análisis incluye los países con al menos un vínculo de colaboración. Fuente: Elaboración propia con R Bibliometrix.

78 países. El análisis determinó que 18 de estos países no colaboraban. El análisis incluyó a los 60 países restantes y la Figura VII presenta los resultados. Los países con mayor número de colaboraciones en la literatura sobre EE son China, EE.UU., España, Reino Unido e Italia. China domina la red de colaboración, seguida de Estados Unidos. Se puede decir que los países con baja participación en la red de colaboración colaboran con países geográficamente cercanos a ellos. Por ejemplo, la colaboración Egipto-EUA, la colaboración Dinamarca-Suiza-Finlandia-Noruega, y la colaboración Perú-Cuba. La colaboración Eslovaquia-República Checa.

Recibido: 17/02/2023 Aceptado: 13/10/2023

Discusión

La eficiencia es un tema de investigación estudiado por muchas disciplinas. La investigación sobre la eficiencia se ha llevado a cabo en muchos ámbitos, como las instituciones educativas, los estudiantes, los programas y las unidades académicas. Ha surgido la necesidad de cartografiar la literatura sobre EE para abordar de forma exhaustiva esta literatura que se ha acumulado a lo largo del tiempo, revisar el campo de la EE y proporcionar una perspectiva a los investigadores. Basándose en esta necesidad, esta investigación pretende abordar la investigación sobre la eficiencia en el campo de la educación desde una perspectiva holística, explorar la estructura intelectual del campo, revelar patrones en la literatura y contribuir al avance del campo de formas nuevas y significativas. Para lograr este objetivo, el estudio buscó respuestas a cuatro preguntas. Éstas son: conocer la bibliografía básica de la educación ambiental, identificar los temas del campo, identificar los temas de tendencia del campo y revelar la interacción social de los países que contribuyen al campo. Para ello, se realizó un análisis bibliométrico de un total de 1.315 publicaciones que abarcaban el periodo comprendido entre 1970 y 2022 utilizando las bases de datos Web of Science y SCOPUS.

En general, la bibliografía sobre EE ha mostrado un crecimiento continuo desde 1970. En concreto, los avances de los últimos 10 años abarcan el 68% de la bibliografía sobre EE que abarca casi 50 años. Se han realizado investigaciones sobre diversos tipos de eficiencias en muchos campos, desde la educación básica hasta las universidades, las unidades académicas y los sistemas educativos de los países. El estudio de De Witte y López-Torres (2017) también lo ha demostrado claramente. En la literatura sobre EE, la razón más fundamental para la mejora continua es la necesidad de hacer el uso más eficaz de los limitados recursos disponibles, como el presupuesto educativo, el personal, la tecnología y la infraestructura. La eficiencia educativa es un importante tema de debate entre los responsables políticos, los profesores y otras partes interesadas. Aparte de la creciente concienciación sobre la eficiencia pública, el aumento del coste de la educación es una de las razones del creciente interés por la eficiencia educativa (De Witte & López-Torres, 2017). Es deseable que todos los niveles educativos, desde la educación básica hasta la educación superior, lleven a cabo sus actividades de la

manera más eficaz. La forma de determinarlo es medir la eficiencia de las unidades de decisión de acuerdo con el propósito. Se puede considerar el crecimiento significativo de la literatura sobre EE, especialmente en los últimos 10 años, como un indicio de que la literatura crecerá para alcanzar los mismos objetivos en el futuro.

Diversas revistas han contribuido a este desarrollo de la literatura sobre EE. La naturaleza interdisciplinar de la investigación en economía de la educación ha contribuido significativamente a la aparición de esta diversidad. Aunque una gran variedad de disciplinas y revistas han contribuido a la literatura sobre EE, la "Revista de Economía de la Educación" ha dominado la literatura. Esta revista lleva publicando muchos artículos de calidad en el campo de la EE desde 1981. Esta revista puede considerarse tanto un recurso crucial para los investigadores que desean publicar trabajos de investigación en el campo de la EE como una importante alternativa en cuanto a oportunidades de publicación.

Si observamos la bibliografía sobre EE desde una perspectiva mundial, el hecho de que China y EE.UU. sean los dos países que más han contribuido a este campo puede atribuirse al mayor número de instituciones de enseñanza superior y, por tanto, al mayor número de investigadores de estos países. Sin embargo, el hecho de que los investigadores de algunos países se especialicen y lleven a cabo investigaciones en el campo de la EE hace que estos países destaquen. Por ejemplo, Agasisti T. (de Italia) resulta ser el autor más prolífico en este campo. Esto contribuyó a que Italia entrara en la categoría de países productivos.

El análisis de palabras clave muestra que DEA es bastante dominante entre los métodos de medición de la eficiencia. Este hallazgo respalda a De Witte y López-Torres (2017), y Villano y Tran (2021). Uno de los métodos utilizados con frecuencia en la medición de la EE es el análisis de frontera estocástica. Sin embargo, hay algunas razones por las que DEA se utiliza más en la medición de la eficiencia. DEA es un método de optimización de la programación matemática. DEA es un método que puede medir la eficiencia de unidades con múltiples entradas y múltiples salidas (Gralka et al., 2019; Mikušová, 2017). El propósito de DEA es determinar la eficiencia del uso de recursos en diferentes organizaciones y las tecnologías disponibles utilizadas y establecer una medida para evaluar los éxitos u oportunidades de conservación de recursos para cada forma de toma de decisiones a la que se asignan los recursos (Charnes et al., 1978). Los métodos no paramétricos de medición de la eficiencia,

Recibido: 17/02/2023 Aceptado: 13/10/2023

como el DEA, han despertado el interés de investigadores de muchos campos de la ciencia, especialmente de los científicos especializados en gestión. Esto se debe a que puede aplicarse para medir la eficiencia técnica de las unidades de toma de decisiones considerando únicamente los insumos y los productos sin ninguna información sobre precios y teniendo en cuenta múltiples insumos y productos (Villano y Tran, 2021). Las instituciones educativas no son organizaciones que intenten alcanzar sus objetivos con un único input y un único output. Por el contrario, funcionan utilizando múltiples inputs y outputs. Las instituciones educativas son organizaciones sin ánimo de lucro que carecen de precios de entrada y salida (Andersson et al., 2017). Las estimaciones de eficiencia basadas en DEA son, por tanto, un enfoque teóricamente sólido para instituciones educativas con múltiples inputs y outputs sin información de precios (Villano & Tran, 2021).

A pesar del predominio metodológico del DEA en la literatura sobre EE, la educación superior tiene un dominio significativo entre los niveles educativos que se han analizado en términos de eficiencia. Tras los drásticos recortes del gasto público en muchos países desarrollados en la década de 1980, la eficiencia en la enseñanza superior se convirtió en un tema de debate particular (Peters, 1992). El rápido aumento del número de estudiantes, unido a estos recortes, suscitó preocupación por el rendimiento de las instituciones de enseñanza superior. Durante este periodo, los gobiernos intentaron satisfacer la creciente demanda de enseñanza superior con los mismos o menores niveles de financiación. Se desarrolló una visión negativa relacionada con la eficiencia debido a la reducción de los recursos asignados por los financiadores públicos (Kupriyanova et al., 2018). Sin embargo, el valor de la eficiencia se restableció con la adopción de nuevos principios de gestión pública en la educación superior (Broucker et al., 2015). El hecho de que la demanda de educación superior siga aumentando en la actualidad, el incremento de los importes de gasto en este sentido y la correcta realización de las misiones y actividades que se esperan de la educación superior hicieron que el tema de la eficiencia en la educación superior siguiera siendo prevalente.

El tema de la eficiencia de la investigación en educación superior, que domina la literatura sobre EE, se ha diversificado con el tiempo. Como indicador de ello se pueden considerar los estudios de eficiencia sobre innovación, transferencia de tecnología y recursos humanos que surgieron en esta investigación. En una economía basada en el conocimiento, las

instituciones de educación superior desempeñan un papel fundamental en el desarrollo económico innovador como productoras y transmisoras de conocimientos (Etzkowitz y Leydesdorff, 1995; Perkmann et al., 2013). Los procesos innovadores se consideran el motor del desarrollo económico (Dunning, 2000). Desde la década de 1970, los países industrializados han puesto en marcha numerosas iniciativas para promover una colaboración más estrecha entre las universidades y la industria, con la esperanza de estimular la innovación y el crecimiento económico. De este modo, el papel de las universidades en materia de innovación se ha ido ampliando gradualmente (Anselin et al., 1997; Feller et al., 2002). Los países con un mayor rendimiento en innovación suelen lograr un mayor crecimiento económico, una mayor competitividad y un mejor nivel de vida (Acs et al., 2002). Los procesos de innovación de las instituciones educativas, y de las universidades en particular, son fundamentales para garantizar un crecimiento económico sostenible (Shi et al., 2020). Este importante vínculo entre las universidades y la innovación ha allanado el camino para la evaluación de la eficiencia innovadora de las universidades en términos regionales, tecnológicos, científicos y empresariales.

En los últimos años, el interés por la eficiencia en la educación ha aumentado en términos prácticos y académicos (Goldstein y Woodhouse, 2000). Paralelamente, la investigación sobre la eficiencia del capital humano ha cobrado importancia en la literatura sobre EE. Una de las razones del surgimiento de la investigación en este ámbito es la creciente importancia de las instituciones educativas en la economía, ya que proporcionan formación intelectual, capital humano de mayor calidad y mayor productividad laboral (Blau, 1996). Son varias las actividades que las universidades pueden realizar para impulsar el desarrollo económico. Entre ellas se encuentran tanto la creación de conocimiento como las relacionadas con la innovación regional a través de la investigación y la transferencia de tecnología (Barra & Zotti, 2017). Los individuos altamente cualificados y bien formados son uno de los principales productos de las universidades. También son el motor del desarrollo económico (Florida et al., 2008). Por lo tanto, se puede decir que los temas del capital humano y la transferencia de tecnología han estado recientemente a la vanguardia del examen de la eficiencia de las universidades.

Este rápido crecimiento de la bibliografía sobre EE, especialmente en los últimos 10 años, puede ser un indicador del crecimiento futuro. Por lo tanto, es posible que en el futuro sea necesario revisar periódicamente

la bibliografía sobre EE. Una revisión más exhaustiva de la bibliografía sobre EE mediante la ampliación de las bases de datos utilizadas en este estudio y la identificación de las tendencias de investigación en este campo puede considerarse un camino a seguir para futuras investigaciones.

La limitación de esta investigación son las bases de datos utilizadas. En este estudio se han utilizado Web of Science y SCOPUS, que no se utilizan juntas en muchos estudios bibliométricos, para cubrir la literatura sobre EE de la forma más amplia posible. Sin embargo, se pueden encontrar estudios sobre literatura de EE en otras bases de datos distintas de estas dos. Estas bases de datos también pueden incluirse en futuras investigaciones. En segundo lugar, aunque se utilizó una estrategia de búsqueda exhaustiva para llegar a todas las publicaciones relevantes, no se pudo llegar a todas las publicaciones relevantes en ambas bases de datos. Esto puede deberse al alcance de la estrategia de búsqueda, a la falta de títulos, resúmenes y palabras clave utilizadas por los autores en sus publicaciones. Los estudios de análisis bibliométrico necesitan datos bibliométricos. WOS y Scopus son dos grandes bases de datos utilizadas con frecuencia para los análisis bibliométricos. En este estudio, a diferencia de muchos otros, se amplió el alcance de la investigación utilizando conjuntamente estas dos grandes bases de datos. Sin embargo, esta investigación no pretende que todos los estudios sobre la eficiencia en la educación estén incluidos en el conjunto de datos. Porque hay muchas bases de datos aparte de estas dos grandes bases de datos. Sin embargo, la mayoría de ellas no proporcionan datos bibliométricos. En consecuencia, aunque el conjunto de datos de esta investigación es amplio, puede haber muchos estudios que no estén incluidos en el análisis.

Cuando se evalúan los resultados de esta investigación en su conjunto, se observa lo siguiente. En primer lugar, existe un interés creciente por la eficacia en la educación. Esta situación tiene por objeto obtener el máximo rendimiento de las inversiones realizadas en educación y de los insumos utilizados. Porque, el uso eficiente de los recursos educativos evita el despilfarro de recursos y maximiza los rendimientos económicos y sociales esperados de la educación. En segundo lugar, esta investigación ha analizado exhaustivamente la literatura sobre la eficiencia educativa con datos a largo plazo. Los resultados de esta investigación ofrecen una visión general del campo a los investigadores o a los investigadores que inician su carrera y trabajan sobre la eficiencia educativa. Los resultados de este estudio proporcionan una visión holística del campo y permiten

a los investigadores especializados en eficiencia educativa examinar la situación actual y trazar futuras líneas de investigación.

Conclusión

A partir de las bases de datos Web of Science y SCOPUS, este artículo ofrece una visión general de la bibliografía sobre EE desde 1970 hasta 2022 a través de un total de 1.315 documentos. También se identifican temas y tendencias en este campo y se aborda la creciente literatura sobre EE desde una perspectiva holística. Los resultados de esta investigación revelan que la bibliografía sobre EE crece a un ritmo del 9% anual. El crecimiento, especialmente en los últimos 10 años, abarca el 68% de toda la bibliografía sobre EE. La publicación clave en este campo es Emrouznejad et al. (2008), que revisa 30 de literatura sobre DEA. La revista clave en este campo es "Revista de Economía de la Educación". El autor que más ha contribuido al campo es Agasisti, T con 38 publicaciones y una media de 28,2 citas por publicación. La institución que más ha contribuido a la literatura sobre EE es la Universidad de Lancaster, en el Reino Unido. De los 78 países que contribuyen al campo, China es el mayor contribuyente y colaborador, seguido de EE.UU. Las palabras que más se repiten son "DEA", "eficiencia" y "enseñanza superior". DEA, SFA, educación superior, productividad, rendimiento, eficiencia técnica, eficiencia de asignación, eficiencia de escala y eficiencia de costes son los temas principales del campo. La evaluación de la eficiencia basada en el examen PISA es un nuevo tema emergente en el campo. En particular, la innovación, la transferencia de tecnología y el capital humano son temas de nicho relacionados con las universidades. Además, la eficiencia de la investigación, la clasificación, la calidad, la eficacia escolar, la competencia, el aprendizaje electrónico y la investigación sobre la eficiencia de la educación europea son algunos de los temas de tendencia.

Referencias bibliográficas

Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review, 22*(1), 89-97. https://doi.org/10.1016/S0272-7757(01)00068-1

Recibido: 17/02/2023 Aceptado: 13/10/2023

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy, 31*(7), 1069-1085. https://doi.org/https://doi.org/10.1016/S0048-7333(01)00184-6

Adamides, E. D., & Karacapilidis, N. (2006). Information technology support for the knowledge and social processes of innovation management. *Technovation, 26*(1), 50-59. https://doi.org/https://doi.org/10.1016/j.technovation.2004.07.019

Afonso, A., & Aubyn, M. S. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic Modelling, 23*(3), 476-491. https://doi.org/10.1016/j.econmod.2006.02.003

Agasisti, T. (2014). The efficiency of public spending on education: an empirical comparison of EU countries. *European Journal of Education, 49*(4), 543-557. https://doi.org/10.1111/ejed.12069

Agasisti, T., Barra, C., & Zotti, R. (2019a). Research, knowledge transfer, and innovation: The effect of Italian universities' efficiency on local economic development 2006-2012. *Journal of Regional Science, 59*(5), 819-849. https://doi.org/10.1111/jors.12427

Agasisti, T., & Dal Bianco, A. (2009). Reforming the university sector: effects on teaching efficiency-evidence from Italy. *Higher Education, 57*(4), 477-498. https://doi.org/10.1007/s10734-008-9157-x

Agasisti, T., Munda, G., & Hippe, R. (2019b). Measuring the efficiency of European education systems by combining data envelopment analysis and multiple-criteria evaluation. *Journal of Productivity Analysis, 51*(2-3), 105-124. https://doi.org/10.1007/s11123-019-00549-6

Agasisti, T., & Zoido, P. (2018). Comparing the efficiency of schools through international benchmarking: Results from an empirical analysis of OECD PISA 2012 data. *Educational Researcher, 47*(6), 352-362. https://doi.org/10.3102/0013189x18777495

Anastasiou, A., Kounetas, K., Mitropoulos, P., & Mitropoulos, Y. (2007). Evaluating efficiency of university academic departments. *Journal of Interdisciplinary Mathematics, 10*(3), 409-431. https://doi.org/10.1080/09720502.2007.10700502

Andersson, C., Antelius, J., Mansson, J., & Sund, K. (2017). Technical efficiency and productivity for higher education institutions in Sweden. *Scandinavian Journal of Educational Research, 61*(2), 205-223. https://doi.org/10.1080/00313831.2015.1120230

Anderson, T. R., Daim, T. U., & Lavoie, F. F. (2007). Measuring the efficiency of university technology transfer. *Technovation, 27*(5), 306-318. https://doi.org/10.1016/j.technovation.2006.10.003

Ang, S., Zheng, R., Wei, F. Q., & Yang, F. (2021). A modified DEA-based approach for selecting preferred benchmarks in social networks. *Journal of the Operational Research Society, 72*(2), 342-353. https://doi.org/10.1080/01605682.2019.1671155

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics, 42*(3), 422-448. https://doi.org/https://doi.org/10.1006/juec.1997.2032

Arbona, A., Giménez, V., López-Estrada, S., & Prior, D. (2022). Efficiency and quality in Colombian education: An application of the metafrontier Malmquist-Luenberger productivity index. *Socio-Economic Planning Sciences, 79*, 101122. https://doi.org/10.1016/j.seps.2021.101122

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959-975. https://doi.org/https://doi.org/10.1016/j.joi.2017.08.007

Aria, M., Cuccurullo, C., D'Aniello, L., Misuraca, M., & Spano, M. (2022). Thematic analysis as a new culturomic tool: The social media coverage on COVID-19 Pandemic in Italy. *Sustainability, 14*(6), 36-43. https://doi.org/https://doi.org/10.3390/su14063643

Auranen, O., & Nieminen, M. (2010). University research funding and publication performance-An international comparison. *Research Policy, 39*(6), 822-834. https://doi.org/10.1016/j.respol.2010.03.003

Avkiran, N. K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences, 35*(1), 57-80. https://doi.org/10.1016/S0038-0121(00)00010-0

Barra, C., & Zotti, R. (2017). Investigating the human capital development-growth nexus: Does the efficiency of universities matter? *International Regional Science Review, 40*(6), 638-678. https://doi.org/10.1177/0160017615626215

Barra, C., & Zotti, R. (2018). The contribution of university, private and public sector resources to Italian regional innovation system (in) efficiency. *The Journal of Technology Transfer, 43*(2), 432-457. https://doi.org/10.1007/s10961-016-9539-7

Bayley, T., Begen, M. A., Rodrigues, F. F., & Barrett, D. (2022). Relative efficiency of radiation treatment centers: an application of data envelopment analysis. *Healthcare, 10*(6). https://doi.org/10.3390/healthcare10061033

Bayraktar, E., Tatoglu, E., & Zaim, S. (2013). Measuring the relative efficiency of quality management practices in Turkish public and private universities. *Journal of the Operational Research Society, 64*(12), 1810-1830. https://doi.org/10.1057/jors.2013.2

Beasley, J. E. (1990). Comparing university departments. *Omega, 18*(2), 171-183. https://doi.org/10.1016/0305-0483(90)90064-G

Beasley, J. E. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society, 46*(4), 441-452. https://doi.org/10.1057/jors.1995.63

Bessent, A. M., & Bessent, E. W. (1980). Determining the comparative efficiency of schools through data envelopment analysis. *Educational Administration Quarterly, 16*(2), 57-75. https://doi.org/10.1177/0013161X8001600207

Blau, F. D. (1996). Symposium on primary and secondary education. *Journal of Economic Perspectives, 10*(4), 3-8.

Broucker, B., De Wit, K., & Leisyte, L. (2015). An evaluation of new public management in higher education. Same rationale, different implementation. EAIR 37th Annual Forum, Krems.

Brzezicki, L. (2020). Efficiency of the university libraries in polish higher education. *E-Mentor*(1), 61-70. https://doi.org/https://doi.org/10.15219/em83.1453

Canal, G. Y., Amado, A. P. G., & Hurtado, M. G. (2015). Research efficiency assessment of Colombian public universities 2003-2012: Data envelopment analysis. *Inge Cuc, 11*(2), 97-108. https://doi.org/10.17981/ingecuc.11.2.2015.10

Chang, T. Y. (2013). Enhancing e-learning management systems to promoting the management efficiency of tourism and hospitality education. *The Anthropologist, 16*(3), 473-485. https://doi.org/10.1080/09720073.2013.11891373

Chang, Y. T., Park, H., Lee, S., & Kim, E. (2018). Have Emission Control Areas (ECAs) harmed port efficiency in Europe? *Transportation Research Part D-Transport and Environment, 58*, 39-53. https://doi.org/10.1016/j.trd.2017.10.018

Chao, S. L., Yu, M. M., & Hsieh, W. F. (2018). Evaluating the efficiency of major container shipping companies: A framework of dynamic network DEA with shared inputs. *Transportation Research Part a-Policy and Practice, 117*, 44-57. https://doi.org/10.1016/j.tra.2018.08.002

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 3*(4), 339-338.

Chen, Z. F., Matousek, R., & Wanke, P. (2018). Chinese bank efficiency during the global financial crisis: A combined approach using satisficing DEA and Support Vector Machines. *North American Journal of Economics and Finance, 43*, 71-86. https://doi.org/10.1016/j.najef.2017.10.003

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics, 5*(1), 146-166. https://doi.org/10.1016/j.joi.2010.10.002

Çetin, M., & Maral, M. (2022). An analysis of research on the efficiency of higher education in Türkiye. *Educational Policy Analysis and Strategic Research, 17*(4), 8-44. https://doi.org/10.29329/epasr.2022.478.1

De Clercq, D., Wen, Z. G., & Fei, F. (2019). Determinants of efficiency in anaerobic bio-waste co-digestion facilities: A data envelopment analysis and gradient boosting approach. *Applied Energy, 253*. https://doi.org/10.1016/j.apenergy.2019.113570

De Witte, K., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society, 68*(4), 339-363. https://doi.org/10.1057/jors.2015.92

Delprato, M., & Antequera, G. (2021). School efficiency in low and middle income countries: An analysis based on PISA for development learning survey. *International Journal of Educational Development, 80*, 1-13. https://doi.org/10.1016/j.ijedudev.2020.102296

Dufrechou, P. A. (2016). The efficiency of public education spending in Latin America: A comparison to high-income countries. *International Journal of Educational Development, 49*, 188-203. https://doi.org/10.1016/j.ijedudev.2016.03.005

Dunning, J. H. (2000). Regions, globalization, and the knowledge economy: the issues stated. In J. H. Dunning (Ed.), *Regions, globalization, and the knowledge-based economy* (pp. 7-41). Oxford Academic.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences, 42*(3), 151-157. https://doi.org/https://doi.org/10.1016/j.seps.2007.07.002

Etzkowitz, H., & Leydesdorff, L. (1995). The Triple Helix-University-industry-government relations: A laboratory for knowledge based economic development. *EASST Review, 14*(1), 14-19.

Feller, I., Ailes, C. P., & Roessner, J. D. (2002). Impacts of research universities on technological innovation in industry: evidence from engineering research centers. *Research Policy, 31*(3), 457-474. https://doi.org/https://doi.org/10.1016/S0048-7333(01)00119-6

Ferro, G., & D'Elia, V. (2020). Higher education efficiency frontier analysis: A review of variables to consider. *Journal on Efficiency and Responsibility in Education and Science, 13*(3), 140-153.

Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development-human capital, the creative class and tolerance. *Journal of Economic Geography, 8*(5), 615-649. https://doi.org/https://doi.org/10.1093/jeg/lbn023

Goldstein, H., & Woodhouse, G. (2000). School effectiveness research and educational policy. *Oxford Review of Education, 26*(3-4), 353-363. https://doi.org/https://doi.org/10.1080/713688547

Gomez-Gallego, J. C., Gomez-Gallego, M., Garcia-Garcia, J. F., & Faura-Martinez, U. (2021). Evaluation of the efficiency of European health systems using fuzzy data envelopment analysis. *Healthcare, 9*(10). https://doi.org/10.3390/healthcare9101270

Gong, L., Liu, Z. Y., Zhang, S. Q., & Jiang, Z. H. (2022). Does open innovation promote innovation efficiency in Chinese universities? *Ieee Transactions on Engineering Management, in press*, 1-14. https://doi.org/10.1109/tem.2021.3140116

Gralka, S., Wohlrabe, K., & Bornmann, L. (2019). How to measure research efficiency in higher education? Research grants vs. publication output. *Journal of Higher Education Policy and Management, 41*(3), 322-341. https://doi.org/10.1080/1360080x.2019.1588492

Hallinger, P. (2021). Tracking the evolution of the knowledge base on problem-based learning: A bibliometric review, 1972-2019. *Interdisciplinary Journal of Problem-Based Learning, 15*(1). https://doi.org/https://doi.org/10.14434/ijpbl.v15i1.28984

Holmgren, J. (2018). The effects of using different output measures in efficiency analysis of public transport operations. *Research in Transportation Business and Management, 28*, 12-22. https://doi.org/10.1016/j.rtbm.2018.02.006

Jiang, J. L., Lee, S. K., & Rah, M. J. (2020). Assessing the research efficiency of Chinese higher education institutions by data envelopment analysis. *Asia Pacific Education Review, 21*(3), 423-440. https://doi.org/10.1007/s12564-020-09634-0

Johnes, J. (2004). Efficiency measurement. In G. Johnes & J. Johnes (Eds.), *The International Handbook on the Economics of Education* (pp. 613–742). Edward Elgar.

Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review, 25*(3), 273-288. https://doi.org/10.1016/j.econedurev.2005.02.005

Johnes, J., Portela, M., & Thanassoulis, E. (2017). Efficiency in education. *J Oper Res Soc, 68*(4), 331-338. https://doi.org/10.1057/s41274-016-0109-z

Kalb, A. (2010). *Public sector efficiency: applications to local governments in Germany*. Gabler Verlag

Kalirajan, K. P., & Shand, R. T. (1999). Frontier production functions and technical efficiency measures. *Journal of Economic Surveys, 13*(2), 149-172.

Kantabutra, S., & Tang, J. C. S. (2010). Efficiency analysis of public universities in Thailand. *Tertiary Education and Management, 16*(1), 15-33. https://doi.org/10.1080/13583881003629798

Kao, C. A., & Hung, H. T. (2008). Efficiency analysis of university departments: An empirical study. *Omega-International Journal of Management Science, 36*(4), 653-664. https://doi.org/10.1016/j.omega.2006.02.003

Khodadadipour, M., Hadi-Vencheh, A., Behzadi, M. H., & Rostamy-Malkhalifeh, M. (2021). Undesirable factors in stochastic DEA cross-efficiency evaluation: An application to thermal power plant energy efficiency. *Economic Analysis and Policy, 69*, 613-628. https://doi.org/10.1016/j.eap.2021.01.013

Kupriyanova, V., Estermann, T., & Sabic, N. (2018). Efficiency of universities: Drivers, enablers and limitations. In *European higher education area: The impact of past and future policies* (pp. 603-618). Springer, Cham.

Lampe, H. W., & Hilgers, D. (2015). Trajectories of efficiency measurement: A bibliometric analysis of DEA and SFA. *European Journal of Operational Research, 240*(1), 1-21. https://doi.org/https://doi.org/10.1016/j.ejor.2014.04.041

Lee, B. L., & Johnes, J. (2022). Using network DEA to inform policy: The case of the teaching quality of higher education in England. *Higher Education Quarterly, 76*(2), 399-421. https://doi.org/doi.org/10.1111/hequ.12307

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management, 45*(2), 175-194. https://doi.org/https://doi.org/10.1177/031289621987767

Liu, F., & Xu, H. (2017). Effects of educational efficiency on national competitiveness based on cross-national data. *Education Sciences, 7*(4), 81. https://doi.org/10.3390/educsci7040081

Lu, K., & Wolfram, D. (2012). Measuring author research relatedness: A comparison of word-based, topic-based, and author cocitation approaches. *Journal of the American Society for Information Science and Technology, 63*(10), 1973-1986. https://doi.org/https://doi.org/10.1002/asi.22628

Maral, M. (2023). Examination of research efficiency of research universities in Türkiye by data envelopment analysis. *International Journal of Educational Management, Advance online publication*. https://doi.org/10.1108/IJEM-04-2023-0194

Mikušová, P. (2017). Measuring the efficiency of the Czech public higher education institutions: an application of DEA. *Journal on Efficiency and Responsibility in Education and Science, 10*(2), 58-63. https://doi.org/https://doi.org/10.7160/eriesj.2017.100204

Moreno-Enguix, M. D., Gomez-Gallego, J. C., & Gallego, M. G. (2018). Analysis and determination the efficiency of the European health systems. *International Journal of Health Planning and Management, 33*(1), 136-154. https://doi.org/10.1002/hpm.2412

Moutinho, V., Madaleno, M., & Macedo, P. (2020). The effect of urban air pollutants in Germany: eco-efficiency analysis through fractional regression models applied after DEA and SFA efficiency predictions. *Sustainable Cities and Society, 59*, 1-11. https://doi.org/10.1016/j.scs.2020.102204

Recibido: 17/02/2023 Aceptado: 13/10/2023

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews, 10*(1), 1-11. https://doi.org/https://doi.org/10.1186/s13643-021-01626-4

Park, Y. S., Lim, S. H., Egilmez, G., & Szmerekovsky, J. (2018). Environmental efficiency assessment of US transport sector: A slack-based data envelopment analysis approach. *Transportation Research Part D-Transport and Environment, 61*, 152-164. https://doi.org/10.1016/j.trd.2016.09.009

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'este, P., Fini, R., Geuna, A., Grimaldi, R., & Hughes, A. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy, 42*(2), 423-442. https://doi.org/https://doi.org/10.1016/j.respol.2012.09.007

Peters, M. (1992). Performance and accountability in 'post-industrial society's: The crisis of British universities. *Studies in Higher Education, 17*(2), 123-139.

Quintano, C., Mazzocchi, P., & Rocca, A. (2021). Evaluation of the eco-efficiency of territorial districts with seaport economic activities. *Utilities Policy, 71*. https://doi.org/10.1016/j.jup.2021.101248

Reichmann, G. (2004). Measuring university library efficiency of using data envelopment analysis. *Libri, 54*(2), 136-146. https://doi.org/10.1515/libr.2004.136

Rhaiem, M. (2017). Measurement and determinants of academic research efficiency: A systematic review of the evidence. *Scientometrics, 110*(2), 581-615.

Saravanan, P., Rajeswari, S., Kumar, J. A., Rajasimman, M., & Rajamohan, N. (2022). Bibliometric analysis and recent trends on MXene research–A comprehensive review. *Chemosphere, 286*, 1-14. https://doi.org/https://doi.org/10.1016/j.chemosphere.2021.131873

Shamohammadi, M., & Oh, D. H. (2019). Measuring the efficiency changes of private universities of Korea: A two-stage network data envelopment analysis. *Technological Forecasting and Social Change, 148*, 1-15. https://doi.org/10.1016/j.techfore.2019.119730

Shi, X., Wu, Y., & Fu, D. (2020). Does University-industry collaboration improve innovation efficiency? Evidence from Chinese firms. *Economic Modelling, 86*, 39-53. https://doi.org/https://doi.org/10.1016/j.econmod.2019.05.004

Sweileh, W. M. (2018). Research trends on human trafficking: A bibliometric analysis using Scopus database. *Globalization and Health, 14*(1), 1-12. https://doi.org/https://doi.org/10.1186/s12992-018-0427-9

Tan, G. K. R. (2006). Efficiency estimates for Singapore manufacturing: new evidence from the Malmquist index. *Applied Economics Letters, 13*(11), 715-721. https://doi.org/10.1080/13504850500401924

Tavares, R. S., Drumond, G. M., Angula Meza, L., & Mexas, M. P. (2018). Efficiency assessment in university libraries. *Transinformacao, 30*(1), 65-79. https://doi.org/10.1590/2318-08892018000100006

Tran, P. P., Kuo, K. C., Lu, W. M., & Kweh, Q. L. (2020). Benchmarking in Vietnam universities: teaching and research and revenue efficiencies. *Asia Pacific Education Review, 21*(2), 197-209. https://doi.org/10.1007/s12564-019-09622-z

Turkan, S., & Ozel, G. (2017). Efficiency of state universities in Turkey During the 2014-2015 academic year and determination of factors affecting efficiency. *Education and Science, 42*(191), 307-322. https://doi.org/10.15390/eb.2017.6980

Thursby, J. G., & Kemp, S. (2002). Growth and productive efficiency of university intellectual property licensing. *Research Policy, 31*(1), 109-124. https://doi.org/10.1016/S0048-7333(00)00160-8

Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly impact* (pp. 285-320). Springer.

Venkadasalam, S., Mohamad, A., & Sifat, I. M. (2020). Operational efficiency of shipping companies Evidence from Malaysia, Singapore, the Philippines, Thailand and Vietnam. *International Journal of Emerging Markets, 15*(5), 875-897. https://doi.org/10.1108/ijoem-07-2019-0493

Vidoli, F., & Ferrara, G. (2015). Analyzing Italian citrus sector by semi-nonparametric frontier efficiency models. *Empirical Economics, 49*(2), 641-658. https://doi.org/10.1007/s00181-014-0879-6

Villano, R. A., & Tran, C. D. T. (2021). Survey on technical efficiency in higher education: A meta-fractional regression analysis. *Pacific Economic Review, 26*(1), 110-135. https://doi.org/10.1111/1468-0106.12310

Visbal-Cadavid, D., Martinez-Gomez, M., & Guijarro, F. (2017). Assessing the efficiency of public universities through DEA. A Case Study. *Sustainability, 9*(8), 1-19. https://doi.org/10.3390/su9081416

Witte, K. D., & López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society, 68*(4), 339-363. https://doi.org/https://doi.org/10.1057/jors.2015.92

Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics, 9*(3), 245-268. https://doi.org/https://doi.org/10.1080/09645290110086126

Xie, H., Zhang, Y., Wu, Z., & Lv, T. (2020). A bibliometric analysis on land degradation: Current status, development, and future directions. *Land, 9*(1), 1-37. https://doi.org/https://doi.org/10.3390/land9010028

Xu, H. Q., Chung, C.-C., & Yu, C. (2022). Visualizing research trends on culture Neuroscience (2008–2021): A bibliometric analysis. *Frontiers in Psychology*, 1-14. https://doi.org/https://doi.org/10.3389/fpsyg.2022.884929

Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods, 18*(3), 429-472. https://doi.org/10.1177/1094428114562629

Apéndice

Estrategia de búsqueda de Web of Science

(TI=(efficienc*) or AK=(efficienc*)) and (TI=(universit* or scholl or college or educat* or tertiary or master or doctor* or bachelor or graduate or phd or academ* or facult* or teach* or student or pupil or department) or AK=(universit* or scholl or college or educat* or tertiary or master or doctor* or bachelor or graduate or phd or academ* or facult* or teach* or student or pupil or department)) AND (AK=("data envelopment" or "DEA" or "stochastic frontier" or "SFA" or Malmquist or "frontier efficienc*" or tobit or "allocative efficienc*" or "technical efficienc*" or "cost efficienc*" or "scale efficienc*" or "economic efficienc*" or "profit efficienc*" or "constant return" or "CCR" or "BCC" or "variable return" or VRS or "decreasing return" or DRS or "non increasing return" or "input oriented" or "output oriented" or "two stage" or "Cobb Douglas" or "Ordinary Least Squares" or "OLS" or "Corrected Ordinary Least Squares" or COLS or "Modified Ordinary Least Squares" or "MOLS" or "Maximum Likelihood Estimation" or "MLE" or "cross sectional" or "panel data" or "fixed effect" or "ran-

dom effect" or "deterministic frontier" or "slack* based" or "free disposal hull" or "free disposable hull" or "FDH" or "dynamic ordinary least square" or "DOLS" or "fully modified ordinary least square" or "FMOLS" or "super efficienc*" or Benchmark* or bootstrap or "decision making unit" or "frontier seperation approach" or "multi stage approach" or input or output or productivit* or "technical change" or "total factor productivity" or "TFP" or "directional distance" or "cost direct" or "cost indirect" or "output distance" or tobit or "order m" or "conditional efficienc*" or "output efficienc*" or "meta frontier" or "joint production" or "hybrid return" or "cost frontier" or translog or "discretionary" or "non discretionary" or probabilistic or "relative efficienc*" or "research efficienc*" or "parametric frontier" or "non parametric frontier" or "parametric approach" or "non parametric approach" or "directional distance function" or "technology distance function") OR AB=("data envelopment" or "DEA" or "stochastic frontier" or "SFA" or Malmquist or "frontier efficienc*" or tobit or "allocative efficienc*" or "technical efficienc*" or "cost efficienc*" or "scale efficienc*" or "economic efficienc*" or "profit efficienc*" or "constant return" or "CCR" or "BCC" or "variable return" or VRS or "decreasing return" or DRS or "non increasing return" or "input oriented" or "output oriented" or "two stage" or "Cobb Douglas" or "Ordinary Least Squares" or "OLS" or "Corrected Ordinary Least Squares" or COLS or "Modified Ordinary Least Squares" or "MOLS" or "Maximum Likelihood Estimation" or "MLE" or "cross sectional" or "panel data" or "fixed effect" or "random effect" or "deterministic frontier" or "slack* based" or "free disposal hull" or "free disposable hull" or "FDH" or "dynamic ordinary least square" or "DOLS" or "fully modified ordinary least square" or "FMOLS" or "super efficienc*" or Benchmark* or bootstrap or "decision making unit" or "frontier seperation approach" or "multi stage approach" or input or output or productivit* or "technical change" or "total factor productivity" or "TFP" or "directional distance" or "cost direct" or "cost indirect" or "output distance" or tobit or "order m" or "conditional efficienc*" or "output efficienc*" or "meta frontier" or "joint production" or "hybrid return" or "cost frontier" or translog or "discretionary" or "non discretionary" or probabilistic or "relative efficienc*" or "research efficienc*" or "parametric frontier" or "non parametric frontier" or "parametric approach" or "non parametric approach" or "directional distance function" or "technology distance function") OR TI=("data envelopment" or "DEA" or "stochastic frontier" or "SFA" or Malmquist or "frontier efficienc*" or tobit or "allocative efficienc*" or "technical efficienc*" or "cost efficienc*" or "scale efficienc*" or

"economic efficienc*" or "profit efficienc*" or "constant return" or "CCR" or "BCC" or "variable return" or VRS or "decreasing return" or DRS or "non increasing return" or "input oriented" or "output oriented" or "two stage" or "Cobb Douglas" or "Ordinary Least Squares" or "OLS" or "Corrected Ordinary Least Squares" or COLS or "Modified Ordinary Least Squares" or "MOLS" or "Maximum Likelihood Estimation" or "MLE" or "cross sectional" or "panel data" or "fixed effect" or "random effect" or "deterministic frontier" or "slack* based" or "free disposal hull" or "free disposable hull" or "FDH" or "dynamic ordinary least square" or "DOLS" or "fully modified ordinary least square" or "FMOLS" or "super efficienc*" or Benchmark* or bootstrap or "decision making unit" or "frontier seperation approach" or "multi stage approach" or input or output or productivit* or "technical change" or "total factor productivity" or "TFP" or "directional distance" or "cost direct" or "cost indirect" or "output distance" or tobit or "order m" or "conditional efficienc*" or "output efficienc*" or "meta frontier" or "joint production" or "hybrid return" or "cost frontier" or translog or "discretionary" or "non discretionary" or probabilistic or "relative efficienc*" or "research efficienc*" or "parametric frontier" or "non parametric frontier" or "parametric approach" or "non parametric approach" or "directional distance function" or "technology distance function"))

Estrategia de búsqueda SCOPUS

(TITLE(efficienc*) OR KEY(efficienc*)) AND (TITLE(universit* or scholl or college or educat* or tertiary or master or doctor* or bachelor or graduate or phd or academ* or facult* or teach* or student or pupil or department) OR KEY(universit* or scholl or college or educat* or tertiary or master or doctor* or bachelor or graduate or phd or academ* or facult* or teach* or student or pupil or department)) AND (TITLE-ABS-KEY("data envelopment" or "DEA" or "stochastic frontier" or "SFA" or Malmquist or "frontier efficienc*" or tobit or "allocative efficienc*" or "technical efficienc*" or "cost efficienc*" or "scale efficienc*" or "economic efficienc*" or "profit efficienc*" or "constant return" or "CCR" or "BCC" or "variable return" or VRS or "decreasing return" or DRS or "non increasing return" or "input oriented" or "output oriented" or "two stage" or "Cobb Douglas" or "Ordinary Least Squares" or "OLS" or "Corrected Ordinary Least Squares" or COLS or "Modified Ordinary

Recibido: 17/02/2023 Aceptado: 13/10/2023

Least Squares" or "MOLS" or "Maximum Likelihood Estimation" or "MLE" or "cross sectional" or "panel data" or "fixed effect" or "random effect" or "deterministic frontier" or "slack* based" or "free disposal hull" or "free disposable hull" or "FDH" or "dynamic ordinary least square" or "DOLS" or "fully modified ordinary least square" or "FMOLS" or "super efficienc*" or Benchmark* or bootstrap or "decision making unit" or "frontier seperation approach" or "multi stage approach" or input or output or productivit* or "technical change" or "total factor productivity" or "TFP" or "directional distance" or "cost direct" or "cost indirect" or "output distance" or tobit or "order m" or "conditional efficienc*" or "output efficienc*" or "meta frontier" or "joint production" or "hybrid return" or "cost frontier" or translog or "discretionary" or "non discretionary" or probabilistic or "relative efficienc*" or "research efficienc*" or "parametric frontier" or "non parametric frontier" or "parametric approach" or "non parametric approach" or "directional distance function" or "technology distance function")) AND (EXCLUDE (SUBJAREA,"MEDI") OR EXCLUDE (SUBJAREA,"PHYS") OR EXCLUDE (SUBJAREA,"ENER") OR EXCLUDE (SUBJAREA,"ENVI") OR EXCLUDE (SUBJAREA,"NURS") OR EXCLUDE (SUBJAREA,"MATE") OR EXCLUDE (SUBJAREA,"HEAL") OR EXCLUDE (SUBJAREA,"BIOC") OR EXCLUDE (SUBJAREA,"AGRI") OR EXCLUDE (SUBJAREA,"DENT") OR EXCLUDE (SUBJAREA,"EART") OR EXCLUDE (SUBJAREA,"NEUR") OR EXCLUDE (SUBJAREA,"CENG") OR EXCLUDE (SUBJAREA,"PHAR") OR EXCLUDE (SUBJAREA,"CHEM") OR EXCLUDE (SUBJAREA,"IMMU") OR EXCLUDE (SUBJAREA,"VETE") OR EXCLUDE (SUBJAREA,"Undefined")) AND (EXCLUDE (PUBYEAR,2023))

Información de contacto: Konaklar District, National Defense University, Beşiktaş, Istanbul, Turkey. E-mail: muammermaral@hotmail.com

La efectividad de la realidad virtual basada en video en la educación artística para el desarrollo de la inteligencia social y emocional en estudiantes de secundaria

The effectiveness of video-based virtual reality in art education for the development of social and emotional intelligence in high school students

https://doi.org/10.4438/1988-592X-RE-2024-406-640

Jing Xu
Yangzhou University

Resumen

El plan de estudios influye en el aspecto intelectual, social y psicoemocional del desarrollo de la personalidad. La comprensión del papel que la edad y la educación tienen en el desarrollo de la personalidad contribuyó al enfoque innovador del aprendizaje social y emocional (SEL). Este estudio compara los cambios en el desarrollo social y emocional en adolescentes mayores (15-19 años) que asistieron a clases de arte e historia del arte con realidad virtual basada en video. Para ello se utilizaron pruebas de psicodiagnóstico, análisis comparativo y métodos de investigación estadística. Las pruebas de inteligencia emocional se basaron en la teoría de los rasgos científicamente probada conocida como los Cinco Grandes. El IPIP-NEO-300 de Johnson se utilizó como una herramienta conveniente para el diagnóstico rápido de un gran grupo de encuestados. Cuatro grupos de estudiantes de último año de varias escuelas secundarias fueron seleccionados al azar para asistir al Programa Académico del Currículo de Arte de 6 semanas, especialmente desarrollado sobre la base del principio de síntesis de las artes. Se realizaron pruebas de psicodiagnóstico al principio y al final del experimento. Los resultados se compararon con los del grupo de control de estudiantes de último año, que no asistieron a clases de arte. El estudio encontró diferencias estadísticamente confiables y significativas en todas las escalas del Cuestionario de Actitud Personal y la prueba IPIP-NEO-300. Demuestra el efecto positivo y estimulante del Programa SEL en el desarrollo de la personalidad,

la autoestima y las habilidades comunicativas. Los presentes hallazgos confirman la hipótesis inicial de que el Programa SEL desarrollado afecta la inteligencia emocional y las habilidades sociales de los estudiantes. Los hallazgos también demuestran la importancia del aprendizaje social y emocional para la orientación profesional, la autodeterminación y la autorrealización (impulsores del éxito).

Palabras clave: programas académicos, adolescentes, Currículo de Arte, aprendiendo, aprendizaje social y emocional.

Abstract

Curriculum influences the intellectual, social and psychoemotional aspect of personality development. The understanding of the role that one's age and education have in personality development contributed to the innovative approach of social and emotional learning (SEL). This study compares changes in social and emotional development in older adolescents (15-19 years old) who attended art and art history classes with video-based virtual reality. For this, psychodiagnostic testing, comparative analysis, and statistical research methods were used. Emotional intelligence tests were based on the scientifically proven trait theory known as the *Big Five*. Johnson's IPIP-NEO-300 was used as a convenient tool for the quick diagnostics of a large group of respondents. Four groups of senior year students from various high schools were randomly selected to attend the 6-week Art Curriculum Academic Program, specially developed based on the principle of synthesis of arts. Psychodiagnostic testing was conducted at the beginning and at the end of the experiment. The results were compared with those for the control group of senior students, who did not attend art classes. The study found statistically reliable and significant differences across all scales of the Self-Attitude Questionnaire and IPIP-NEO-300 test. It proves the positive, stimulating effect of the SEL Program on personality development, self-esteem and communication skills. The present findings confirm the initial hypothesis that the developed SEL Program affects the emotional intelligence and social skills of students. The findings also demonstrate the importance of social and emotional learning for career guidance, self-determination and self-fulfillment (success drivers).

Keywords: academic programs, adolescents, Art Curriculum, learning, social and emotional learning.

Introducción

El problema de la adolescencia, que según la clasificación actual de la OMS abarca la edad de 10 a 19 años, ha sido objeto de numerosas

investigaciones en psicología educativa (Altwaijri et al., 2020). Durante este tiempo, una persona atraviesa dos períodos de crisis:

- la crisis relacionada con la pubertad a la edad de 13 años; y
- la crisis de 17 años, la cual es una crisis de la edad adulta, cuando se completa la educación escolar, se asume un nuevo papel adulto y ocurre la autodeterminación profesional y la planificación de una vida futura independiente.

Este es un período único: una edad crítica que afecta todas las áreas del desarrollo mental humano, incluidos los aspectos motivacionales y cognitivos, cuando el desarrollo y la consolidación de los hábitos sociales y emocionales se convierten en la clave para asegurar la salud y el bienestar futuros, el moldeo social exitoso del individuo (Altwaijri et al., 2020). Por lo tanto, los programas académicos para adolescentes y jóvenes son de particular interés porque influyen en el moldeo de la personalidad y el desarrollo intelectual y psicoemocional. La comprensión de la importancia de esta edad en la vida de un individuo y el papel determinante de la educación en el moldeo de la personalidad contribuyó al surgimiento del enfoque innovador del aprendizaje social y emocional (SEL). Este enfoque se originó allá por la década de 1970 (Weissberg, 2019) y no ha perdido su importancia desde entonces, volviéndose cada vez más popular en los últimos años. A pesar de las publicaciones periódicas de artículos de investigación científica y monografías sobre cuestiones de SEL, muchos aspectos prácticos siguen siendo contradictorios y poco desarrollados. Esto determinó el propósito del estudio: comparar los cambios en el desarrollo social y emocional de adolescentes mayores (de 15 a 19 años según la clasificación de la OMS) que asistieron a clases de conceptos básicos de arte e historia del arte, utilizando realidad virtual basada en video.

Revisión bibliográfica

La Realidad Virtual (RV) es una tecnología que proporciona inmersión y experiencia a través de modelado tridimensional (Wu et al., 2023). Es una representación esférica interactiva que brinda a los participantes una sensación de presencia (Abich et al., 2021). Aunque los investigadores señalan el impacto positivo de la realidad virtual en el rendimiento

académico de los estudiantes, existen ciertos problemas asociados con su uso en el aula (Lindquist et al., 2020). La RV puede afectar de manera diferente el estado fisiológico de los estudiantes y causar mareos en algunos casos individuales (Farič et al., 2019). Al mismo tiempo, el alto costo y la complejidad de la implementación complican el uso de la RV (Lin et al., 2021). Además, para los profesores es difícil de desarrollar materiales didácticos en la RV, ya que no tienen los conocimientos requeridos (Yu, 2023).

En los últimos años, se han desarrollado una serie de sistemas y recursos de la realidad virtual (SV-RV) basados en vídeo esférico. Los sistemas pueden cumplir con los requisitos del entorno de aprendizaje (Wu et al., 2023). Son más asequibles y solo requieren el uso de una cámara panorámica para crear vídeos (Lindquist et al., 2020). El vídeo inmersivo permite al espectador ver libremente vídeos de 360 grados, controlando el contenido y los ángulos de visión deseados (Farič et al., 2019). Además, esta tecnología facilita solucionar el problema de la excesiva dependencia de la realidad virtual del modelado 3D. En comparación con las tecnologías de realidad virtual basadas en modelos 3D, la realidad virtual inmersiva basada en vídeo esférico (SV-IVR) ahorra una cantidad significativa de tiempo. Debido al progreso tecnológico, el desarrollo de aplicaciones móviles y la popularidad de las cámaras de 360 grados, es posible crear imágenes o vídeos esféricos de 360 grados (Wu et al., 2023). Por lo tanto, la SV-IVR ofrece oportunidades excepcionales para la educación (Yu, 2023).

El aprendizaje social y emocional (SEL) se enfoca en:

- adquirir los conocimientos, actitudes y habilidades necesarias para ser consciente de las propias emociones y sentimientos y poder gestionarlos;
- mostrar empatía por otras personas, desarrollar y mantener buenas relaciones con ellas;
- establecer y alcanzar objetivos significativos y tomar decisiones responsables (Thierry et al., 2022).

SEL crea oportunidades y contribuye a la solución efectiva de tareas y problemas cotidianos (Weissberg, 2019). La combinación de SEL con otros enfoques similares mejora las competencias internas, interpersonales y cognitivas de los estudiantes (Murano et al., 2020). SEL proporciona a los estudiantes competencias sociales y emocionales como la autoconciencia,

Recibido: 14/03/2023 Aceptado: 09/02/2024

la autogestión, la conciencia social, las habilidades para relacionarse y la toma responsable de decisiones (Humphrey et al., 2020). Las funciones preventivas de SEL fortalecen la salud física y mental, contribuyen a la calidad de vida y al éxito en el mercado laboral (Jones et al., 2019).

El desarrollo social, emocional y cognitivo están interconectados en el cerebro y el comportamiento humano, lo que afecta los resultados del aprendizaje, el estilo de vida y el éxito social. Como resultado, cuestiones relacionadas se han convertido en el foco de atención de las autoridades educativas y sanitarias, los políticos, las organizaciones no gubernamentales y los medios de comunicación (Jones et al., 2019).

Rasgos sociales y emocionales de los adolescentes:

- se evalúan mediante herramientas de psicodiagnóstico para determinar la autoestima, la perseverancia y la autoeficacia (Primi et al., 2019);
- dependen de factores de personalidad y afectan aspectos socioeconómicos (Primi et al., 2019).

Probado por Boon-Falleur et al. (2020), correlacionándose con el comportamiento, el plan de estudios de arte debe satisfacer los intereses y las necesidades socioemocionales de los estudiantes (Eddy et al., 2021), ya que representan un componente crucial de la educación de los adolescentes (Hoffmann et al., 2020) y afectan el desarrollo de la personalidad y el bienestar (Green et al., 2021).

El aprendizaje social y emocional se fundamenta en la idea de que las artes son fundamentales para el desarrollo humano y que cada estudiante tiene derecho a desarrollar el pensamiento creativo

(Ewing, 2020). La naturaleza específica de las formas artísticas en un contexto educativo se traduce en la inclusión de elementos de juego y otros medios que influyen en el desarrollo emocional (Ewing, 2020), el pensamiento crítico y creativo (Ewing, 2020). La participación en actividades artísticas tiene un efecto positivo en la confianza en uno mismo, la autoestima, la construcción de relaciones y el sentido de pertenencia, cualidades asociadas con la resiliencia y la salud mental. Los programas SEL normalizan la autoestima de los estudiantes, mejorando su confianza en uno mismos (Gueldner et al., 2020). Dichos programas tienen un impacto positivo en la comprensión lectora, el aprendizaje de matemáticas y los conceptos básicos de la investigación (Cipriano et al., 2023). Sin embargo, si bien reconocen las importantes oportunidades

ofrecidas por los programas SEL, los investigadores (Omasta et al., 2021) advierten contra su popularización basada en los beneficios externos percibidos. El desarrollo de las habilidades sociales y emocionales de los adolescentes mejora los logros personales. Sin embargo, evaluación de tales habilidades incluye una serie de desafíos visionarios y metodológicos que requieren enfoques metodológicos innovadores (Guo et al., 2023). Se puede usar SEL como un método para mejorar la efectividad general y la percepción positiva del programa académico (Mussey, 2019).

La división en programas académicos generales y programas SEL es puramente teórica porque esencialmente todo el proceso de enseñanza tiene un enfoque social y emocional (Frey et al., 2019). La creación de entornos de aprendizaje en los que los estudiantes colaboren con los instructores para adquirir habilidades sociales y emocionales se considera una alternativa positiva al modelo tradicional de disciplina escolar (Voith et al., 2020). El modelo integral SEL ayuda a:

- desarrollar el sentido de identidad de los estudiantes y su confianza en sus capacidades, y desarrollar habilidades de comunicación social (Mahoney et al., 2021);
- satisfacer las necesidades emocionales, mejorando así la eficacia de la enseñanza (Van Tassel-Baska et al., 2021).

La importancia del desarrollo social y emocional de los estudiantes contribuye al uso de dicho desarrollo como una medida del aprendizaje junto con medidas del rendimiento académico (Mahoney et al., 2021).

Objetivos de investigación

- Analizar los datos de investigación disponibles sobre las cuestiones del aprendizaje social y emocional (SEL), la evaluación de sus métodos y los programas académicos del plan de estudios de arte;
- desarrollar un programa académico de arte e historia del arte y determinar su eficacia en el desarrollo de la inteligencia social y emocional de los estudiantes de secundaria (de 16 a 17 años);
- procesar los resultados, usando herramientas estadísticas; discutir la importancia de dichos resultados y sacar conclusiones apropiadas.

Recibido: 14/03/2023 Aceptado: 09/02/2024

La hipótesis de la investigación declara que existe una correlación directa entre la matriculación en un programa académico de arte e historia del arte y la inteligencia social y emocional de los estudiantes de secundaria.

Hay varias razones para elegir una lección de artes visuales para realizar una investigación. La lección de artes visuales tiene una amplia gama de elementos visuales, como colores, formas, texturas, etc. Puede abarcar diversas técnicas y estilos artísticos, creando un entorno de aprendizaje multidimensional. Las lecciones de arte promueven la expresión creativa de los estudiantes al permitirles expresar sus pensamientos, sentimientos e ideas a través del proceso creativo. El uso de la realidad virtual puede ampliar aún más las posibilidades de expresión creativa. El programa académico seleccionado en historia del arte se basa en el principio de síntesis del arte. Incluye elementos de diversas disciplinas artísticas, proporcionando un enfoque de aprendizaje único e integrado. Las artes visuales no solo promueven el desarrollo cognitivo de los estudiantes, sino que también afectan su desarrollo social y emocional. Por lo tanto, la realidad virtual puede permitir una conexión emocional aún más profunda con el arte. El uso de la realidad virtual aporta elementos de innovación y modernidad al proceso educativo, brindando a los estudiantes la oportunidad de interactuar con las últimas tecnologías.

Materiales y métodos

Se usó un diseño de investigación cuasi experimental, lo que implicó dividir a los participantes en un grupo de control y un grupo experimental. Se eligió Historia del Arte como disciplina a enseñar. Los métodos de investigación aplicados en el estudio incluyeron pruebas de psicodiagnóstico, un análisis comparativo y métodos de investigación estadística.

Participantes

El estudio seleccionó al azar a dos grupos de estudiantes de secundaria de diferentes escuelas secundarias para realizar un programa académico de Historia del Arte de seis semanas de duración diseñado sobre el principio

de síntesis del arte. Había 86 estudiantes en el grupo experimental (42 chicos y 44 chicas) y 89 estudiantes en el grupo de control (43 chicos y 46 chicas). La edad promedio de los encuestados en el grupo experimental fue de 15,7 años y la edad promedio de los encuestados en el grupo de control fue de 15,8 años. En el grupo experimental se aplicó el sistema de enseñanza de SV-IVR, mientras que en el grupo de control se aplicó el método tradicional de enseñanza mediante tecnología. Antes del inicio de la enseñanza experimental, cada participante del grupo experimental consiguió un acceso a Google Cardboard. Los participantes no tuvieron problemas de visión ni de audición. Los participantes del estudio representaban diferentes estratos socioeconómicos de una región china. En particular, en el programa participaron estudiantes de secundaria de ciudades y pueblos. La composición etnocultural de la muestra de la investigación también fue diversa, abarcando diferentes grupos culturales y étnicos de la región. Dada la diversidad de los participantes, el contexto sociocultural pudo haber influido en su percepción de la realidad virtual durante el programa académico. Este hecho puede permitir generalizar los resultados del estudio para diversos grupos sociales y etnoculturales de esta región.

Proceso de aprendizaje

El sistema de aprendizaje de SV-IVR se basa en un proceso de aprendizaje sistemático y cíclico. Según el contenido de aprendizaje, los estudiantes entran la ubicación de cada módulo para experimentar y aprender a usar el sistema. Las indicaciones de la voz de fondo facilitan el proceso. En un tiempo limitado, los estudiantes deberán responder preguntas relacionadas con sus conocimientos. Si responden correctamente a todas las preguntas, podrán pasar al siguiente módulo; de lo contrario, tendrán que volver a estudiar el módulo reprobado. Este proceso de aprendizaje cíclico permite a los estudiantes aprender cada módulo en profundidad. El sistema también incluye diversas actividades interactivas. El profesor hace preguntas rápidas que invitan a la reflexión, y luego ayuda a los estudiantes a encontrar respuestas durante el proceso de aprendizaje. Los estudiantes deben cargar sus respuestas en una base de datos interna. La expectativa era que a través de esta interacción se alentaría a los estudiantes a ser proactivos e involucrados, brindándoles múltiples niveles de retroalimentación y reflexión y guiándolos hacia un pensamiento

Recibido: 14/03/2023 Aceptado: 09/02/2024

de alto nivel. Las sesiones de 40 minutos se realizaron tres veces por semana durante un mes y medio. Durante el experimento, los estudiantes tuvieron la oportunidad de participar tanto individualmente como en grupo, dependiendo del diseño de las clases. Algunas tareas y actividades requirieron interacción individual con la realidad virtual. Por lo tanto, cada participante podría concentrarse lo máximo posible en su experiencia y estudiar el material. Por otro lado, algunas tareas estimularon la interacción colectiva, permitiendo a los estudiantes trabajar en grupos. Estas tareas contribuyeron a las habilidades de comunicación, la comprensión mutua y la cooperación, que también son aspectos importantes del desarrollo social. El papel del profesor en este contexto fue principalmente facilitador. El profesor actuó como guía, ayudó a los estudiantes a trabajar con la realidad virtual, brindó explicaciones y contribuyó a la discusión de los resultados. El profesor también supervisó el progreso general del grupo y el desarrollo individual de cada estudiante. Como resultado, fue posible adaptar el proceso educativo a las necesidades de cada participante.

Programa

El programa académico presuponía el estudio de la historia del Arte y las Artes Aplicadas a través de una experiencia de inmersión. Se enfocaba en la revelación de períodos y estilos clave en el mundo del arte y el desarrollo de habilidades creativas. El programa usaba excursiones virtuales interactivas a museos y exposiciones, y reconstrucciones virtuales de obras maestras artísticas. También empleaba la RV para dibujo y modelado interactivos e involucraba a artistas contemporáneos para dar clases maestras virtuales. Había varias visitas virtuales de 360 grados por talleres, estudios y galerías de arte. Los estudiantes realizaron tareas creativas en un entorno virtual, usando cortometrajes sobre la vida y obra de grandes artistas como el contenido de vídeo. Las exposiciones virtuales tenían como objetivo desarrollar habilidades en crítica y análisis del arte, mientras que los laboratorios virtuales interactivos facilitaron experimentos con formas y colores. Los videos permitieron a los estudiantes dominar técnicas artísticas, estimulando su pensamiento creativo y su inmersión emocional. Los objetivos del programa eran desarrollar habilidades de colaboración, discusión y cooperación grupal, así como ampliar la percepción y expresión de emociones a través del proceso creativo.

Herramientas de investigación

El estudio psicodiagnóstico se realizó dos veces: antes y después del experimento. Se utilizaron métodos de psicodiagnóstico válidos: el Cuestionario de Actitud Personal (Psylab, 2021) y el test IPIP-NEO-120 (Sleep et al., 2021), el cual es una versión moderna del clásico test de los *Cinco Grandes*. La distinción entre el autoconcepto (conocimiento o ideas sobre sí mismo, una evaluación de la expresión de ciertas características) y la actitud hacia sí mismo se fundamenta en la base teórica del Cuestionario de Actitud Personal. El cuestionario revela tres niveles de autoestima, que varían en el grado de generalización:

- actitud global hacia uno mismo;
- actitud hacia uno mismo, la cual se divide en autoestima, orgullo, interés propio y expectativas sobre uno mismo;
- nivel de acciones específicas (preparación para tales acciones) que le conciernen a uno mismo.

Las pruebas de inteligencia emocional evalúan la capacidad de una persona para reconocer y comprender tanto sus propias emociones como las de los demás, en las cuales se basan sus interacciones con los demás. Dichas pruebas se basan en una teoría de rasgos con base científica conocida como los *Cinco Grandes*. Esta técnica usa un análisis factorial de descripciones verbales de rasgos humanos, que describe cómo las personas se perciben entre sí y cómo los individuos se adaptan a su entorno, teniendo en cuenta sus rasgos biológicos. Como su nombre lo indica, el modelo de los *Cinco Grandes* examina cinco rasgos de la personalidad (disposiciones) comunes y relativamente independientes:

- extroversión;
- amabilidad (amigabilidad, capacidad de negociación);
- escrupulosidad;
- neuroticismo (siendo la estabilidad emocional su contrario);
- apertura a la experiencia (intelecto).

En 2014, J.A. Johnson elaboró el cuestionario IPIP-NEO-120, una versión abreviada del cuestionario IPIP-NEO-300 creado en 1999 por L. Goldberg, autor del modelo *Cinco Grandes*, el cual consta de cinco escalas que incluyen 30 aspectos del modelo. El estudio actual usó este cuestionario

Recibido: 14/03/2023 Aceptado: 09/02/2024

que consta de 120 preguntas, como una herramienta más conveniente para diagnosticar rápidamente a un grupo grande de encuestados.

Procesamiento de datos estadísticos

Los resultados de la investigación se procesaron, usando el software Excel y una calculadora en línea.

Cuestiones éticas

Este estudio cumple con la Convención sobre los Derechos del Niño, los principios del consentimiento informado y otros principios de la bioética. Dada la edad de los encuestados, el permiso para realizar el experimento se obtuvo no solo del Comité de Ética, sino también de los instructores (que ayudaron en la implementación del programa y monitorearon el cumplimiento de los principios de bioética) y los padres.

Resultados

Primero, los autores realizaron un estudio comparativo de un grupo experimental y un grupo de control, usando el Cuestionario de Actitud Personal. Los resultados diferían ligeramente entre los grupos, pero fueron comparables y, para algunas escalas, no hubo diferencias estadísticamente significativas. En todas las escalas prevalecían las puntuaciones medias, lo cual es lógico, ya que reflejan la norma. Sin embargo, aunque sus puntuaciones se mantuvieron dentro de la norma, ambos grupos obtuvieron las puntuaciones más bajas en la escala de *Actitudes Esperadas de los Demás*. Esta observación sugiere que algunos encuestados tenían problemas de comunicación y expectativas de una actitud adversa por parte de los demás, lo que podría resultar tanto de rasgos de la personalidad como de experiencias adversas.

El grupo experimental y el grupo de control mostraron resultados iniciales comparables en IPIP-NEO-120. No se observaron diferencias estadísticamente significativas en altruismo, modestia, autoconciencia y aventurerismo. Aquí

también prevalecían los valores medios, pero se observaron puntuaciones relativamente más bajas en actividad, colaboración y, en particular, en autoeficacia y autodisciplina. También vale mencionar las puntuaciones moderadamente altas en imaginación y ensoñación (que resultaron estar poco desarrolladas en algunos de los adolescentes estudiados). Podrían indicar un rasgo de personalidad, pero la imaginación debería ser más viva en la infancia y la adolescencia. Por lo tanto, los autores tienden a conceder mayor importancia a los factores sociales y a la falta de atención al desarrollo creativo del individuo. Las siguientes tablas presentan los resultados de las pruebas psicodiagnósticas del grupo experimental realizadas antes y después del programa académico experimental en arte: Tabla I – Indicadores del Cuestionario de Actitud Personal y Tabla 2 – los *Cinco Grandes*.

TABLA I. Indicadores de Actitudes Personales de los encuestados del grupo experimental antes (A) y después (B) del programa académico en arte

Escala	prueba t de Student	p
S (Actitud personal global)	3.713	>0.05
I (Autoestima)	3.531	>0.05
II (Autoafección)	4.743	>0.05
III (Actitudes esperadas de los demás)	4.238	>0.05
IV (Interés propio)	4.238	>0.05
1 (Confianza en sí mismo)	6.576	>0.05
2 (Actitudes por parte de los demás)	4.743	>0.05
3 (Autoaceptación)	4.743	>0.05
4 (Autoguía, autoconsistencia)	6.069	>0.05
5 (Autoculpabilidad)	3.333	>0.05
6 (Interés propio)	2.953	>0.05
7 (Autocomprensión)	5.405	>0.05

Fuente: Elaboración propia.

Recibido: 14/03/2023 Aceptado: 09/02/2024

TABLA II. Resultados de IPIP-NEO-120 entre los encuestados del grupo experimental antes (A) y después (B) del programa académico en arte

Escala	Subescala	prueba t de Student	p
Extroversión	amigabilidad	6.748	<0.05
	Gregarismo	4.065	<0.05
	Asertividad	4.238	<0.05
	Nivel de Actividad	5.240	<0.05
	En búsqueda de emociones	3.713	<0.05
	Alegría	5.571	<0.05
amabilidad	Confianza	1.092	>0.05
	Moralidad	5.240	<0.05
	Altruismo	3.531	<0.05
	Colaboración	5.736	<0.05
	Modestia	2.184	<0.05
	Simpatía	5.075	<0.05
Escrupulosidad	Autoeficacia	7.275	<0.05
	Ordenación	5.736	<0.05
	Sentido de deber	3.537	<0.05
	Esfuerzos hacia los logros	4.238	<0.05
	Autodisciplina	5.902	<0.05
	cautela	5.405	<0.05
Neuroticismo	Ansiedad	4.910	<0.05
	Ira	6.406	<0.05
	Depresión	5.736	<0.05
	Timidez	7.275	<0.05
	Inmoderación	3.152	<0.05
	Vulnerabilidad	7.455	<0.05
Apertura a la experiencia	Imaginación	7.638	<0.05
	intereses artísticos	4.408	<0.05
	Emocionalidad	6.748	<0.05
	Aventurerismo	4.910	<0.05
	Intelecto	2.953	<0.05
	Liberalismo	0.445	>0.05

Fuente: Elaboración propia.

En estos estudios, los autores utilizaron la prueba t de Student para poblaciones relacionadas, lo cual permitió comparar los resultados antes y después de la introducción del programa académico especial. Los autores obtuvieron diferencias estadísticamente fiables y significativas en todas las escalas de la *Actitud Personal*. Los encuestados del grupo experimental tuvieron un aumento estadísticamente significativo en el respeto a sí mismos (la puntuación alta fue alcanzada por el 20,9% de los encuestados inicialmente y luego por el 29,1% de los encuestados en el estudio de seguimiento); interés propio (la puntuación alta fue mostrada por el 18,6% de los encuestados al inicio del estudio y luego por el 32,6% de los encuestados después de completar el programa académico en Arte);la autoafección (la puntuación alta fue en el 16,3% de los encuestados inicialmente y luego en el 29,1% de los encuestados); y los indicadores de la escala de *Actitudes Esperadas de los Demás* – tal aumento parece ser particularmente importante (la puntuación alta fue alcanzada por el 22,1% de los encuestados frente al 11,6% inicialmente, mientras que el número de examinados con puntuaciones bajas cambió del 12,8% al 5,8%). Los indicadores de autoaceptación y autoculpabilidad se normalizaron.

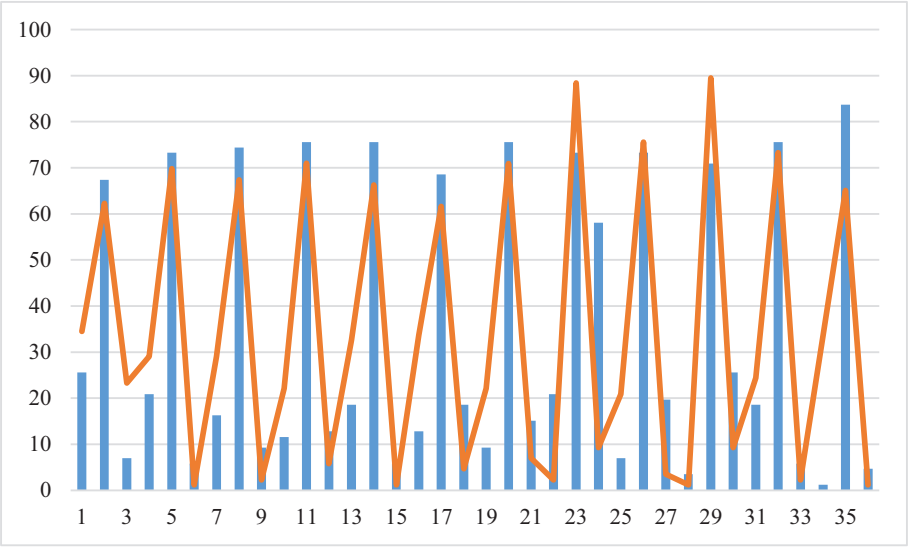
Después del experimento, el porcentaje de los encuestados con puntuaciones altas en la escala IPIP-NEO-120 aumentó significativamente. Hubo mejoras en los siguientes rasgos. *Escala de Extroversión:* amigabilidad (del 24,4 al 43,0%); gregarismo (del 18,6 al 26,9%); asertividad (de 14,0 a 18,6%); nivel de actividad (del 19,8 al 30,2%). *Escala de amabilidad*: confianza (del 27,9 al 37,2%); moralidad (del 17,4 al 32,6%); altruismo (del 19,8 al 27,9%); colaboración (del 15,1 al 32.6%); y simpatía (del 22,1 al 37,3%). *Escala de escrupulosidad*: autoeficacia (del 11,6 al 31,4%); ordenación (del 9,3 al 18,6%); sentido de deber (del 8,1 al 20,9%); esfuerzos hacia los logros (del 23,3 al 32,6%); autodisciplina (del 10,5% al 24,4%). *Escala de Apertura a la Experiencia*: imaginación (del 7,0 al 32,6%); intereses artísticos (del 36,0 al 51,2%); emocionalidad (del 14,0 al 39,5%); intelecto (del 33,7 al 40,7%). El porcentaje de los encuestados con puntuaciones promedio de modestia y cautela aumentó del 61,6% y 64,0% al 77,9% y 80,2%, respectivamente. La misma tendencia se observó en el caso del aventurerismo. No hubo diferencias estadísticamente significativas en las puntuaciones de liberalismo.

Se logró la normalización en la *Escala de Neuroticismo*, como lo demuestra un aumento en el porcentaje de puntuaciones promedio en ansiedad (del 74,4 al 84,8%), ira (del 77,9 al 80,2%), timidez (del 57,0

Recibido: 14/03/2023 Aceptado: 09/02/2024

al 79,1%) e inmoderación (del 76,7 al 84,9%). Los cambios positivos en las puntuaciones de depresión (del 2,3% al 20,9%) y vulnerabilidad (del 7,0 al 27,9%) se manifiestan como un aumento en el porcentaje de los encuestados con puntuaciones bajas. La Figura 1 muestra cambios en algunos indicadores en el grupo experimental. No se observaron cambios estadísticamente significativos entre las puntuaciones de seguimiento y de referencia en el grupo de control. Por lo tanto, no hay razón para presentar sus resultados.

FIGURA I. Resultados iniciales (azul) y de seguimiento (rojo) de la metodología del *Cuestionario de Actitud Personal*

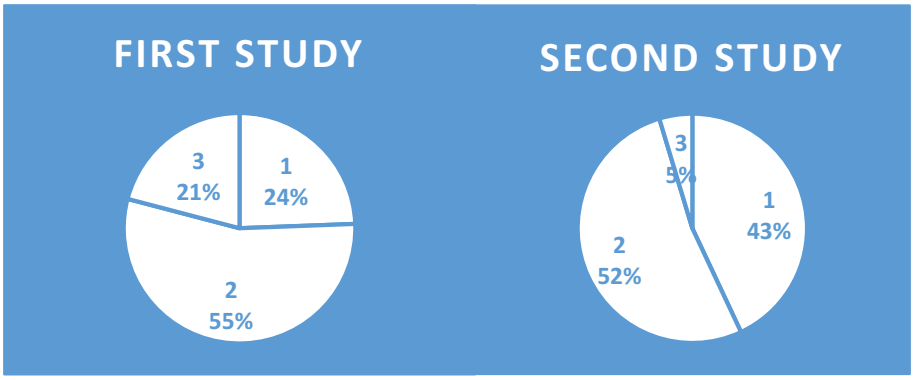


Fuente: Elaboración propia.

Este diagrama ilustra que el rendimiento general, tanto antes como después del experimento, se encuentra en concordancia con la norma de edad. Sin embargo, los autores observaron una mejora visible en varias escalas (es decir, un mayor porcentaje de los encuestados con puntuaciones más altas), lo que sugiere un efecto positivo y estimulante del programa propuesto sobre el desarrollo de la personalidad, la autoestima y las habilidades de comunicación.

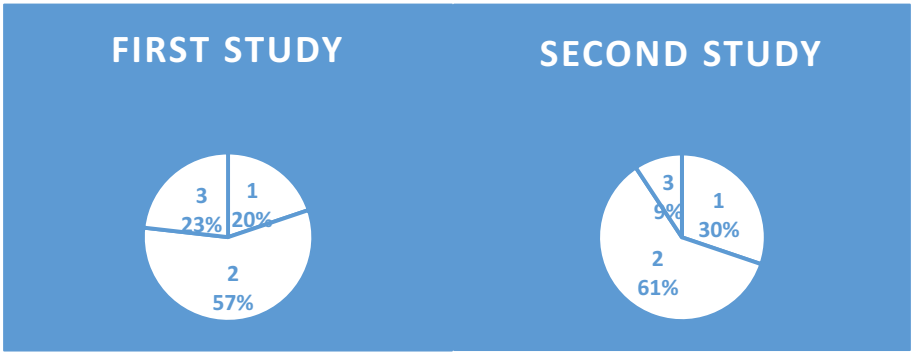
Los siguientes diagramas ilustran los cambios en algunas subescalas del IPIP-NEO-120.

FIGURA II. El porcentaje de encuestados con puntuaciones de amigabilidad alta (1), media (2) y baja (3)



Fuente: Elaboración propia.

FIGURA III. El porcentaje de encuestados con puntuaciones de nivel de actividad alto (1), medio (2) y bajo (3)



Fuente: Elaboración propia.

El formato del artículo no permite hacer una representación gráfica de los cambios para todas las escalas y subescalas del IPIP-NEO-120. Sin embargo, las Figuras de 2 a 4 ofrecen una visión holística del efecto positivo del programa académico en arte en los aspectos sociales y emocionales de las personalidades de los encuestados.

FIGURA IV. El porcentaje de encuestados con puntuaciones altas (1), medias (2) y bajas (3) en intereses artísticos



Fuente: Elaboración propia.

Discusión

Se observa el efecto positivo del programa académico propuesto en arte sobre la actitud personal de los adolescentes en general y sobre los componentes de dicha actitud personal. Esto se refiere a los encuestados esencialmente sanos que mostraron una normalidad promedio en todos los resultados de las pruebas durante el estudio inicial. Las diferencias entre el grupo experimental y el grupo de control fueron mínimas; no se observaron diferencias estadísticas para las escalas individuales. Sin embargo, algunas desviaciones de los indicadores de sus niveles medios indicaron problemas de autoestima y comunicación de los examinados individuales, las cuales pueden interpretarse tanto como rasgos de la personalidad como defectos de educación o un impacto adverso de la sociedad. Estas desviaciones de los niveles medios se observaron en ambas pruebas, entre cuyos indicadores se observó una correlación positiva pronunciada (con un coeficiente de correlación = 0,83). El análisis de los indicadores de las escalas y subescalas del Cuestionario de Actitud Personal revela una serie de puntos clave. La escala de Autorrespeto muestra la capacidad de los examinados para controlar su vida, ser coherentes, creer en sus capacidades y comprenderse a sí mismos. En general, en el estudio inicial se observaron puntuaciones medias predominantemente buenas (el 73,3% de los encuestados) e incluso puntuaciones altas (el 20,9% de los encuestados) en esta escala. Sin embargo, el 5,8% de los

encuestados obtuvo una puntuación baja (menos de 50 puntos), lo que sugiere la falta de independencia y confianza en sí mismos. Las clases de arte permitieron a estos adolescentes descubrir su creatividad personal (Guo et al., 2023), lo que les dio confianza y determinación, como se demuestra en el mayor número de los encuestados que mejoraron sus puntuaciones en esta escala en el estudio de seguimiento, cuando el 29,1% de los examinados demostró un nivel alto, y solo el 1,2% tuvo un nivel bajo.

La escala de *Autoafección* muestra la actitud del examinado (amistosa u hostil) hacia sí mismo. Como se encontró en el estudio inicial, casi el 10% de los encuestados tienden a tratarse a sí mismos de una manera adversa al no aceptarse o culparse a sí mismos.

Además, las respuestas en términos de las escalas adicionales complementaron y aclararon las respuestas en las escalas principales, mostrando una correlación directa con ellas (con un coeficiente de correlación de 0,96). Las actitudes adversas en esta escala indicaron la percepción de las propias deficiencias en detrimento de sus propios lados fuertes, baja autoestima y disposición a culparse a sí mismo. Las actitudes personales adversas pueden expresarse en las correspondientes respuestas emocionales adversas (Boon-Falleur et al., 2020). Las clases de arte mejoraron la situación en esta escala, cambiando la actitud hacia sí mismo, la autoaceptación y la autoestima.

Las clases de arte permitieron a los adolescentes probar su creatividad y descubrir talentos ocultos, habilidades para actividades en las que antes no habían pensado (Ewing, 2020). Esto aumentó el interés propio, así como el interés por los propios pensamientos y sentimientos, y generó confianza en que uno también podría ser interesante para otras personas.

Este cambio se muestra en la escala *Actitudes Esperadas de los Demás*, donde el 7,0% de los encuestados cambió sus expectativas de negativas a positivas. Alrededor del 22,1% de los encuestados mostraron una puntuación alta durante el estudio de seguimiento, mientras que dicha puntuación se observó solo en el 11,6% de los casos durante el estudio inicial. Es probable que la expectativa adversa inicial se debiera a una experiencia pasada fallida, a una crítica ofensiva o a la propia insatisfacción con los intentos de realizar actividades artísticas (Jagers et al., 2019). El programa SEL promovió la autoexpresión y la autoafirmación de los estudiantes, y la creatividad colectiva ayudó a desarrollar habilidades comunicativas, lo cual es consistente con la literatura de

Recibido: 14/03/2023 Aceptado: 09/02/2024

investigación (Sprenger, 2020). Por lo tanto, la presente investigación sobre la metodología del Cuestionario de Actitud Personal demuestra el efecto positivo del programa académico en arte en el desarrollo social y emocional de los estudiantes.

Para respaldar esta declaración, los autores utilizaron el cuestionario IPIP-NEO-120 que evalúa los rasgos de la personalidad de una persona en comparación con los rasgos de la personalidad de otras personas. La escala de extroversión muestra la tendencia de una persona a comunicarse y acctuar, mientras que las puntuaciones bajas en esta escala son más típicas de los introvertidos (personas que prefieren el silencio, la soledad y la desvinculación social). Los adolescentes se caracterizan por una mayor comunicación, un comportamiento gregario e incluso en el mundo digital actual se comunican mucho, no solo a través de reuniones físicas, sino también en las redes sociales, donde son muy activos. Por lo tanto, puntuaciones bajas en esta escala pueden ser una manifestación de introversión como un rasgo personal y pueden sugerir distanciamiento social debido a experiencias adversas y neurotización de la personalidad. Esta suposición está respaldada por un aumento general en las puntuaciones, tanto en la escala de Extroversión (y sus subescalas) como en la normalización de las puntuaciones en la escala de Neuroticismo y la subescala de Depresión. La puntuación media de extroversión de la mayoría de los examinados sugiere que sus emociones están equilibradas. Las puntuaciones bajas en la subescala de amigabilidad pueden estar relacionadas con la moderación emocional como un rasgo de personalidad, o pueden reflejar hostilidad hacia el mundo que los rodea. Las puntuaciones de esta subescala se complementan con la subescala de sociabilidad, cuyas puntuaciones bajas indican una preferencia por la privacidad (Kolyadenko et al., 2008).

El aumento en la puntuación en la escala de extroversión no significa que la metodología de educación artística propuesta busque que todos los adolescentes sean activos socialmente a pesar de sus rasgos innatos de la personalidad. Más bien, se enfoca en superar las barreras de comunicación interna, el miedo y la desconfianza del mundo que nos rodea, promoviendo una comunicación eficaz en el proceso de una actividad significativa y creativa. Los adolescentes de hoy suelen pasar demasiado tiempo frente a la computadora, experimentando problemas para controlar su tiempo y atención. Los adolescentes suelen asociar la actividad con desafíos emocionantes y vivir al límite, mientras que la

pasividad se asocia con la pereza y los problemas de autoorganización. El aumento en la actividad en el presente estudio se distinguió por estar asociado con el fortalecimiento de la actividad creativa, lo cual constituyó un incentivo para la comunicación y el aprendizaje, el desarrollo personal y el autoconocimiento.

La subescala de actividad se correlaciona con la subescala de vigor que mide el estado de ánimo y los sentimientos positivos, mientras que las puntuaciones bajas en esta escala señalan manifestaciones de neuroticismo. El aumento en la puntuación en esta escala se explica por el hecho de que la creación de objetos de arte contribuyó a las emociones positivas, la alegría por los logros, el surgimiento del entusiasmo y los planes optimistas para el futuro.

Las puntuaciones elevadas en la subescala de amabilidad sugieren una armonización de las relaciones sociales de los encuestados y el desarrollo de sus habilidades de colaboración. Las puntuaciones bajas en la escala de confianza indican que otras personas son percibidas como potencialmente peligrosas, egoístas y engañosas. Las clases propuestas permitieron a los adolescentes superar estos miedos sociales en gran medida, por eso el 37,2% de los encuestados en el estudio de seguimiento mostraron puntuaciones altas frente al 27,9% en el estudio inicial, y el número de los encuestados con puntuaciones bajas de confianza cayó del 12,8% al 5,8%. La desconfianza suele ir acompañada de una mala comprensión de los fundamentos de la moralidad, aceptando la mentira como un patrón de conducta justificable. Por lo tanto, fue muy satisfactorio encontrar un efecto positivo de las clases de arte en el aumento en las puntuaciones de la subescala de moralidad, donde inicialmente se observaron puntuaciones altas en el 17,4% de los encuestados, y en el 32,6% de los encuestados durante el estudio de seguimiento.

La creatividad colaborativa en las clases promovió no solo la autocomprensión, sino también la empatía y, como resultado, aumentó el altruismo y la colaboración (Lang et al., 2020). Por lo tanto, el programa académico en arte contribuyó a las habilidades de socialización positiva de los adolescentes estudiados, la capacidad de involucrarse en actividades creativas colaborativas, que les serán útiles en su futura vida adulta.

La actividad intelectual presupone el establecimiento de objetivos, la capacidad de organizar y planificar formas de lograrlos (Sleep et al., 2021). Una puntuación media en la escala de escrupulosidad sugiere una organización y un autocontrol suficientes.

Recibido: 14/03/2023 Aceptado: 09/02/2024

La autodisciplina refleja el componente volitivo de la personalidad, el cual se manifiesta en la capacidad de superar el deseo de posponer las cosas, de no distraerse y de persistir en tareas difíciles o desagradables hasta que estas se lleven a cabo.

La normalización de estos indicadores es un factor predictivo porque muestra el efecto de las clases de arte en el desarrollo de una habilidad vital (Frey et al., 2019) que será útil en futuras carreras y contribuirá al éxito personal.

Las personas con mayor neuroticismo tienen más probabilidades de interpretar las situaciones normales como amenazantes y las frustraciones menores como desesperadamente difíciles (Sleep et al., 2021). La normalización de las puntuaciones en esta escala parece ser particularmente importante con respecto a la subescala de depresión, porque en la adolescencia incluso una respuesta depresiva contextual puede desencadenar un comportamiento impulsivo, agresivo o autoagresivo. La normalización de la puntuación de neuroticismo sugiere el desarrollo de la capacidad de afrontar situaciones estresantes y frustrantes, controlar las emociones y encontrar soluciones significativas a los problemas que surgen.

La imaginación es otro rasgo en el que concentrarse. Los autores asumieron que los estudiantes deberían tener una mejor imaginación en comparación con los adultos, pero recibieron en su mayoría cifras promedio (el 73,3% de los encuestados) y cifras por debajo del promedio (Kraehe, 2020). Los autores tienden a explicar esta situación, así como un aumento significativo en el indicador debido a las clases de arte (un nivel alto de esta subescala en el estudio de seguimiento se observó en el 32,6% de los encuestados del grupo experimental) por falta de los factores que estimulan la ensoñación que es esencialmente la base del pensamiento innovador, una de las competencias educativas clave en la sociedad moderna. El enfoque de la enseñanza en la obtención de soluciones claras y listas para usar, y el miedo a ser malinterpretados, al verbalizar las propias ideas resultan ser factores disuasorios para la autoexpresión creativa de los estudiantes. En este contexto, las clases de arte ayudan a desarrollar un pensamiento libre de clichés, a dar un impulso a la búsqueda creativa, armonizando así el desarrollo personal de los adolescentes. El estudio inicial reveló que el 36,0% de los encuestados del grupo experimental y el 37,1% de los encuestados del grupo de control habían estado interesados en el arte antes, pero a menudo lo trataban de forma

distante, como consumidores. Ahora, gracias al programa académico especial, los estudiantes del grupo experimental se han familiarizado con el arte como sus creadores, y esto ha aumentado su interés por este tema (el 51,2% de los encuestados en el estudio de seguimiento). En el transcurso del experimento, muchos estudiantes superaron sus miedos e incomodidades en situaciones sociales, se volvieron más equilibrados y abiertos a experiencias positivas, lo que sugiere un cambio en sus estilos cognitivos hacia el desarrollo de la creatividad y el sentido de la belleza. Por lo tanto, este estudio confirma la hipótesis original del efecto del programa SEL propuesto en la inteligencia emocional y las habilidades sociales de los estudiantes, y demuestra la importancia del aprendizaje social y emocional a través del programa académico en arte para la orientación profesional, la autodeterminación y la autorrealización de los estudiantes como factores de éxito personal futuro.

Limitaciones

El estudio incluyó solo a unas pocas escuelas secundarias y una pequeña muestra de encuestados. Los resultados pueden considerarse representativos de todos los estudiantes de último año de una región determinada, pero otros grupos de edad y grupos etnoculturales pueden revelar resultados diferentes. Además, el artículo no comparó la efectividad de diferentes programas académicos y métodos de enseñanza de SEL. Este tema puede ser el foco de investigaciones futuras.

Conclusiones

El modelo integral SEL ayuda a desarrollar el sentido de sí mismos y la confianza en los estudiantes, instilar la capacidad de superar dificultades y promover habilidades de comunicación social, tales como el trabajo en equipo y el intercambio de experiencias. Las mejoras estadísticamente confiables y significativas ($p < 0,05$) en todas las escalas de Actitud Propia e IPIP-NEO-120 apoyan esta declaración. El programa académico de arte contribuyó al desarrollo de habilidades de socialización positiva y mejoró la capacidad de los estudiantes para participar en actividades creativas colaborativas, lo cual será útil en su vida adulta. La normalización de

Recibido: 14/03/2023 Aceptado: 09/02/2024

puntuaciones en la escala de Neuroticismo confirma el impacto social y emocional positivo del programa académico en arte. Esto parece ser particularmente importante con respecto a la subescala de Depresión. En la adolescencia, incluso una respuesta depresiva contextual puede desencadenar un comportamiento impulsivo, agresivo o autoagresivo. La normalización sugiere que los estudiantes desarrollaron la capacidad de afrontar situaciones estresantes y frustrantes, controlar las emociones y encontrar soluciones significativas a los problemas que surgen.

El enfoque en la obtención de soluciones claras y listas para usar en clase y el miedo a ser malinterpretado al verbalizar las propias ideas resultan ser factores disuasorios para la autoexpresión creativa. Las clases de arte ayudan a desarrollar un pensamiento libre de clichés, a impulsar la búsqueda creativa, armonizando así el desarrollo personal de los adolescentes. En el transcurso del experimento, muchos estudiantes superaron sus miedos e incomodidades en situaciones sociales, se volvieron más equilibrados y abiertos a experiencias positivas, lo cual sugiere un cambio en sus estilos cognitivos hacia el desarrollo de la creatividad y el sentido de la belleza. Las puntuaciones aumentadas en la escala de inteligencia señalan una capacidad cognitiva simbólica y, dependiendo de la capacidad particular de cada individuo, pueden tomar la forma de razonamiento matemático, lógico o geométrico, uso artístico y metafórico del lenguaje, composición o interpretación musical, o una de las muchas artes visuales o escénicas.

El estudio del programa SEL basado en la realidad virtual en artes visuales para adolescentes mayores reveló varias implicaciones y ventajas potenciales para la educación y otras áreas. Los resultados del estudio indican un efecto positivo del programa SEL en la inteligencia emocional y las habilidades sociales de los estudiantes. Este beneficio puede ser un aspecto clave en la implementación de este y otros programas similares para estudiantes de secundaria. El estudio confirma que el programa influye en la autoestima y las habilidades comunicativas de los estudiantes. Por lo tanto, puede convertirse en la base para mejorar el desarrollo personal y preparar la comunicación futura en la sociedad actual. Los resultados indican la importancia del aprendizaje social y emocional para la orientación profesional, la autodeterminación y la autorrealización. Este hecho puede ser una motivación para implementar este tipo de programas a gran escala en el sistema educativo. El éxito de la implementación de la realidad virtual en la educación en artes visuales

puede convertirse en un modelo para utilizar enfoques similares en la enseñanza de otras materias dirigidas al desarrollo social y emocional. Los métodos innovadores de aprendizaje social y emocional se pueden aplicar a otros grupos de edad, dependiendo de las necesidades y personalidades individuales de los estudiantes. El estudio también puede abrir oportunidades para el uso de la realidad virtual para el desarrollo social y emocional en otros campos, tales como la psicología, el desarrollo personal y el aprendizaje corporativo.

Referencias bibliográficas

Abich, J., Parker, J., Murphy, J. S., & Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality*, *25*(4), 919-933. https://doi.org/10.1007/s10055-020-00498-8

Altwaijri, Y. A., Al-Subaie, A. S., Al-Habeeb, A., Bilal, L., Al-Desouki, M., Aradati, M., King, A. J., Sampson, N. A., & Kessler, R. C. (2020). Lifetime prevalence and age-of-onset distributions of mental disorders in the Saudi National Mental Health Survey. *International Journal of Methods in Psychiatric Research*, *29*(3), e1836. https://doi.org/10.1002/mpr.1836

Boon-Falleur, M., Bouguen, A., Charpentier, A., Algan, Y., Huillery, E., & Chevallier, C. (2020). *Measuring socio-emotional skills in schools: simple questionnaires outperform behavioral tasks.* PsyArXiv PPR325803. https://doi.org/10.31234/osf.io/w7pr6

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C., & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development*, *94*(5), 1181-1204. https://doi.org/10.1111/cdev.13968

Eddy, M., Blatt-Gross, C., Edgar, S. N., Gohr, A., Halverson, E., Humphreys, K., & Smolin, L. (2021). Local-level implementation of Social Emotional Learning in arts education: Moving the heart through the arts. *Arts Education Policy Review*, *122*(3), 193-204. https://doi.org/10.1080/10632913.2020.1788681

Recibido: 14/03/2023 Aceptado: 09/02/2024

Ewing, R. (2020). Introduction. In *Transforming the Curriculum through the Arts* (pp. 1-3). Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-52797-6_1

Farič, N., Potts, H. W., Hon, A., Smith, L., Newby, K., Steptoe, A., & Fisher, A. (2019). What players of virtual reality exercise games want: Thematic analysis of web-based reviews. *Journal of Medical Internet Research, 21*, e13833. https://doi.org/10.2196/13833

Frey, N., Fisher, D., & Smith, D. (2019). *All learning is social and emotional: Helping students develop essential skills for the classroom and beyond*. ASCD.

Green, A. L., Ferrante, S., Boaz, T. L., Kutash, K., & Wheeldon-Reece, B. (2021). Social and emotional learning during early adolescence: Effectiveness of a classroom-based SEL program for middle school students. *Psychology in the Schools*, *58*(6), 1056-1069. https://doi.org/10.1002/pits.22487

Gueldner, B. A., Feuerborn, L. L., & Merrell, K. W. (2020). *Social and emotional learning in the classroom: Promoting mental health and academic success*. Guilford Publications.

Guo, J., Tang, X., Marsh, H. W., Parker, P., Basarkod, G., Sahdra, B., Ranta, M., & Salmela-Aro, K. (2023). The roles of social–emotional skills in students' academic and life success: A multi-informant and multicohort perspective. *Journal of Personality and Social Psychology, 124*(5), 1079-1110. https://doi.org/10.1037/pspp0000426

Hoffmann, J. D., Brackett, M. A., Bailey, C. S., & Willner, C. J. (2020). Teaching emotion regulation in schools: Translating research into practice with the RULER approach to social and emotional learning. *Emotion*, *20*(1), 105-109. https://doi.org/10.1037/emo0000649

Humphrey, N., Lendrum, A., Wigelsworth, M., & Greenberg, M. T. (2020). *Social and emotional learning*. Routledge.

Jagers, R. J., Rivas-Drake, D., & Williams, B. (2019). Transformative social and emotional learning (SEL): Toward SEL in service of educational equity and excellence. *Educational Psychologist, 54*(3), 162-184. https://doi.org/10.1080/00461520.2019.1623032

Jones, S. M., McGarrah, M. W., & Kahn, J. (2019). Social and emotional learning: A principled science of human development in context. *Educational Psychologist, 54*(3), 129-143. https://doi.org/10.1080/00461520.2019.1625776

Kolyadenko, N. V., Zdorovenko, N. V., Yaroslavskaya, S. M., Zhivago, K. S., & Kostyuk, Y. A. (2008). *Method of synthetic psychotherapy: Patent 37555, Ukraine, IPC 7 A61H3 / 00; applicant and patent owner National Medical University them. OO Bogomolets. - No. 200811197, application 16.09.2008; publ. 25.11.2008; Bull. No 22 (Ua).* https://iprop-ua.com/inv/1djl8k5m/

Kraehe, A. M. (2020). Dreading, pivoting, and arting: The future of art curriculum in a post-pandemic world. *Art Education, 73*(4), 4-7. https://doi.org/10.1080/00043125.2020.1774320

Lang, S. N., Jeon, L., Sproat, E. B., Brothers, B. E., & Buettner, C. K. (2020). Social emotional learning for teachers (SELF-T): A short-term, online intervention to increase early childhood educators' resilience. *Early Education and Development*, *31*(7), 1112-1132. https://doi.org/10.1080/10409289.2020.1749820

Lin, H. C. S., Yu, S. J., Sun, J. C. Y., & Jong, M. S. Y. (2021). Engaging university students in a library guide through wearable spherical video-based virtual reality: Effects on situational interest and cognitive load. *Interactive Learning Environments, 29*(8), 1272-1287. https://doi.org/10.1080/10494820.2019.1624579

Lindquist, M., Maxim, B., Proctor, J., & Dolins, F. (2020). The effect of audio fidelity and virtual reality on the perception of virtual greenspace. *Landscape and Urban Planning, 202*, 103884. https://doi.org/10.1016/j.landurbplan.2020.103884

Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128-1142. https://doi.org/10.1037/amp0000701

Murano, D., Sawyer, J. E., & Lipnevich, A. A. (2020). A meta-analytic review of preschool social and emotional learning interventions. *Review of Educational Research*, *90*(2), 227-263. https://doi.org/10.3102/0034654320914743

Mussey, S. (2019). *Mindfulness in the Classroom: Mindful Principles for Social and Emotional Learning (1st ed.)*. Routledge. https://doi.org/10.4324/9781003236665

Omasta, M., Graham, M., Milling, S. L., Murray, E., Jensen, A. P., & Siebert, J. J. (2021). Social emotional learning and the national core arts standards: A cross-disciplinary analysis of policy and practices. *Arts Education Policy Review, 122*(3), 158-170. https://doi.org/10.1080/10632913.2020.1773366

Primi, R., Santos, D. D. D., Hauck, N., Fruyt, F. D., & John, O. P. (2019). Mapping self-report questionnaires for socio-emotional characteristics: What do they measure? *Estudos de Psicologia (Campinas), 36*, e180138. https://doi.org/10.1590/1982-0275201936e180138

Psylab (2021). *Self-attitude questionnaire*. https://psylab.info/

Sleep, C. E., Lynam, D. R., & Miller, J. D. (2021). A comparison of the validity of very brief measures of the Big Five/Five-Factor Model of personality. *Assessment*, *28*(3), 739-758. https://doi.org/10.1177/1073191120939160

Sprenger, M. (2020). *Social-emotional learning and the brain: Strategies to help your students thrive*. Ascd.

Thierry, K., Vincent, R., Norris, K., & Dawson, O. (2022). Improving school readiness for a social emotional learning curriculum: Case study of a school-mental health agency Partnership. *Journal of Research in Childhood Education*, *36*(3), 483-505. https://doi.org/10.1080/02568543.2021.1960938

Van Tassel-Baska, J., Cross, T. L., & Olenchak, F. R. (2021). *Social-emotional curriculum with gifted and talented students.* Routledge.

Voith, L. A., Yoon, S., Topitzes, J., & Brondino, M. J. (2020). A feasibility study of a school-based social emotional learning program: Informing program development and evaluation. *Child and Adolescent Social Work Journal*, *37*, 329-342. https://doi.org/10.1007/s10560-019-00634-7

Weissberg, R. P. (2019). Promoting the social and emotional learning of millions of school children. *Perspectives on Psychological Science, 14*(1), 65-69. https://doi.org/10.1177/1745691618817756

Wu, W.-L., Hsu, Y, Yang, Q.-F., Chen, J.-J., & Jong, M. S.-Y. (2023). Effects of the self-regulated strategy within the context of spherical video-based virtual reality on students' learning performances in an art history class. *Interactive Learning Environments, 31(4)*, 2244-2267. https://doi.org/10.1080/10494820.2021.1878231

Yu, Z. (2023). A meta-analysis of the effect of virtual reality technology use in education. *Interactive Learning Environments, 31*(8), 4956-4976. https://doi.org/10.1080/10494820.2021.1989466

Información de contacto: Jing Xu. College of Music, Yangzhou University, Yangzhou, China. E-mail: jingxu68@gmx.com

Recibido: 14/03/2023 Aceptado: 09/02/2024

Problemas aritméticos verbales en los libros castellano-leoneses y andaluces

Arithmetic word problems in Castilian-Leonesian and Andalusian textbooks

https://doi.org/10.4438/1988-592X-RE-2024-406-641

Santiago Vicente Martín
https://orcid.org/0000-0002-2072-8133
Universidad de Salamanca

Marta Ramos Baz
https://orcid.org/0000-0002-9643-6495
Universidad de Salamanca

Mª. Mercedes Rodríguez Sánchez
https://orcid.org/0000-0002-1020-6681
Universidad de Salamanca

Beatriz Sánchez-Barbero
https://orcid.org/0000-0001-9118-2118
Universidad de Salamanca

Rosario Sánchez Fernández
https://orcid.org/0000-0002-7984-1894
Universidad de Salamanca

Resumen

Los resultados de evaluaciones del rendimiento en matemáticas como TIMSS, en las que la resolución de problemas es la piedra angular, muestran que el rendimiento los alumnos de Castilla y León es significativamente superior al

Recibido: 14/03/2023 Aceptado: 09/02/2024

de los alumnos de Andalucía. Estas diferencias parecen deberse a algunos aspectos señalados por el propio TIMSS (como el Índice Social y Económico), aunque también podrían deberse en algún grado a determinados elementos del proceso de enseñanza-aprendizaje, como el libro de texto, puesto que se usan de manera intensiva en las aulas. De hecho, los libros utilizados en países con alumnos más competentes en matemáticas incluyen más problemas, y problemas más variados y difíciles que los libros de otros países. El objetivo del presente trabajo es determinar si los libros de matemáticas de Primaria que se utilizan en Castilla y León son diferentes a los que se utilizan en Andalucía, en relación con el aprendizaje de la resolución de problemas aritméticos verbales. Para ello, en un primer estudio, se determina qué libros se utilizan con más frecuencia en cada comunidad autónoma, y en un segundo estudio, se analizan libros de 3º de Primaria de los proyectos editoriales más utilizados en cada comunidad autónoma en relación con la cantidad de tareas destinadas a resolver problemas, y la variedad y nivel de dificultad semántico-matemática de esos problemas. Los resultados indicaron que, aunque los libros más frecuentes en ambas comunidades pertenecen a proyectos editoriales diferentes, los libros más utilizados en Castilla y León y Andalucía eran similares respecto a las variables analizadas, encontrándose como única diferencia una mayor cantidad de problemas en los libros castellano-leoneses. Estos resultados contrastan con los obtenidos en comparaciones internacionales de libros de texto, y apuntan a una mayor influencia de otras variables en las diferencias de rendimiento de los alumnos de ambas comunidades.

Palabras clave: Libro de texto, enseñanza primaria, resolución de problemas, aritmética, análisis comparativo, análisis de contenido.

Abstract

The results of mathematics achievement assessments such as TIMSS, in which problem solving is the cornerstone, show that the performance of students in Castile and Leon is significantly higher than that of students in Andalusia. These differences seem to be due to some aspects pointed out by TIMSS (such as the Social and Economic index), although they could also be due to some degree to certain elements of the teaching-learning process, such as the textbook, since they are used intensively in the classroom. In fact, the books used in countries with more proficient students in mathematics include more problems, and more varied and difficult problems than books from other countries. The aim of the present work is to determine whether the primary mathematics textbooks used in Castile and Leon are different from those used in Andalusia, in relation to the learning of verbal arithmetic problem solving. For this purpose, in a first study, we determined which books are most frequently used in each autonomous community, and in a second study we analyzed 3rd grade primary school books

Recibido: 14/03/2023 Aceptado: 09/02/2024

from the most widely used textbooks in each autonomous community in relation to the number of tasks aimed at solving problems, and the variety and level of semantic-mathematical difficulty of these problems. The results indicated that, although the most frequently used books in both communities belong to different publishing projects, the most used books in Castile and Leon and Andalusia were very similar with respect to the variables analyzed, the only difference being a greater number of problems in the Castilian-Leonese books. These results contrast with those obtained in international comparisons of textbooks and point to a greater influence of other variables in the differences in student performance in both autonomous communities.

Keywords: textbook, primary education, problem solving, arithmetic, comparative analysis, content analysis.

Introducción

La resolución de problemas es la piedra angular del currículo de matemáticas de la mayoría de los sistemas educativos del mundo (Philpot et al., 2021) y de los marcos teóricos de las evaluaciones internacionales de rendimiento en matemáticas (Mullis et al., 2020). En este sentido, el marco teórico del Trends in International Mathematics and Science Study (TIMSS) define la resolución de problemas como la habilidad de resolver una variedad amplia de situaciones, aplicando los conocimientos matemáticos necesarios para realizar razonamientos de diferentes tipos y niveles de dificultad (Lindquist et al., 2017).

Los resultados de las evaluaciones internacionales muestran que no todos los países ni sus diferentes regiones están siendo igual de eficaces al desarrollar en sus alumnos la habilidad de resolver problemas. Para comprender el origen de esas diferencias, TIMSS proporciona datos sobre algunas variables educativas y socio-culturales: entorno familiar, composición y recursos de la escuela, clima escolar o la seguridad y disciplina del centro escolar (Mullis et al., 2020). Además de esas variables, otros elementos más concretos de la práctica educativa, como los libros de texto, influyen en el aprendizaje de los alumnos (Fagginger et al., 2016; Sievert et al., 2019, 2021).

Considerando lo anterior, en el presente trabajo se pretende comprobar si en dos comunidades autónomas españolas, cuyos alumnos han mostrado niveles de competencia matemática muy dispares, se

utilizan libros diferentes, tanto en relación con su presencia en las aulas, como con la cantidad y variedad de problemas aritméticos verbales (en adelante, PAV) que contienen.

Diferencias en el Rendimiento en Resolución de Problemas

Uno de los objetivos de los programas como TIMSS es evaluar el éxito que alcanzan diferentes sistemas educativos al desarrollar distintas habilidades matemáticas en los alumnos, entre las que destaca la resolución de problemas (Lindquist et al., 2017). TIMSS establece diferentes niveles de rendimiento en función de las puntuaciones de los alumnos, que implican distintos grados de habilidad para resolver problemas. Los alumnos con un nivel bajo tendrían dificultades para resolver cualquier problema; los alumnos con un nivel intermedio únicamente podrían resolver problemas rutinarios; los alumnos con nivel alto resolverían problemas de dos pasos que requieran la comprensión conceptual de los números enteros; finalmente, los alumnos con un nivel avanzado podrían resolver problemas complejos de múltiples pasos que requieran comprensión y razonamiento (Mullis et al., 2020).

Los resultados de TIMSS (2016) en España revelaron que, en Castilla y León (en adelante, CyL), el porcentaje de alumnos de 4º de Primaria con niveles alto (33%) y avanzado (7%) es significativamente mayor que en Andalucía (18% y 2%, respectivamente). Asimismo, se encontraron diferencias significativas en todos los dominios evaluados, tanto de contenido (números, formas y medidas geométricas y representaciones de datos), como cognitivos (conocer, razonar y aplicar).

Las diferencias de rendimiento entre alumnos de distintos países y regiones parecen deberse a múltiples causas: implicación de las familias en la educación escolar, importancia de la educación para la sociedad, presupuesto dedicado a la educación, o formación del profesorado (Rao et al., 2010). Los datos del propio TIMSS y de otros informes, como PISA, indican una fuerte correlación entre el rendimiento en matemáticas y el Índice Social, Económico y Cultural de las familias (ISEC, ver Instituto Nacional de Evaluación Educativa, I.N.E.E., 2023). Otros estudios atribuyeron las diferencias entre regiones de un mismo país a otros factores (creencias epistemológicas de los profesores, expectativas laborales de los alumnos tras su escolaridad, o determinadas variables

Recibido: 14/03/2023 Aceptado: 09/02/2024

instruccionales, como el tiempo y tipo de instrucción que reciben los alumnos, ver Hippe et al., 2018). En España, las diferencias entre alumnos castellano-leoneses y andaluces se han atribuido a factores generales y difíciles de modificar, como un mayor ISEC de las familias, una mayor equidad del sistema educativo de CyL o a una mayor ansiedad ante las matemáticas de los alumnos andaluces (I.N.E.E., 2023). No obstante, el análisis de determinadas variables instruccionales más cercanas a la práctica educativa podría ayudar a entender estas diferencias. Una de esas variables instruccionales podría ser el libro de texto y el tipo de actividades y problemas que en ellos se encuentren. En este sentido, los PAV cobran una especial relevancia.

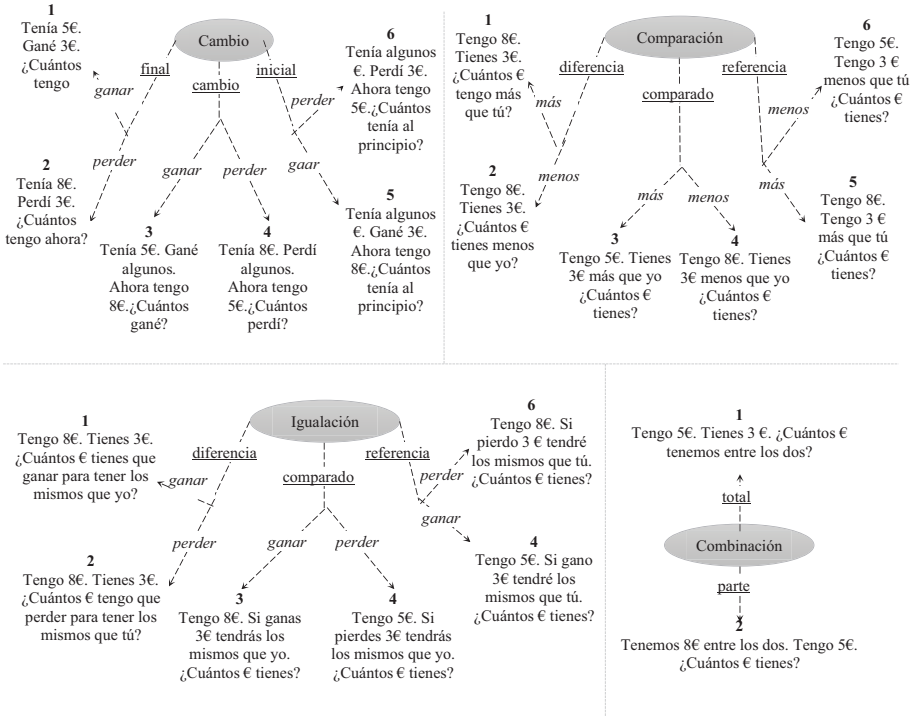
Qué son los Problemas Aritméticos Verbales

Los PAV son descripciones verbales de situaciones en las que se plantea una o varias preguntas cuyas respuestas pueden obtenerse aplicando operaciones aritméticas a los datos numéricos del problema (Verschaffel et al., 2020). Los PAV pueden entrañar diferentes niveles de dificultad para los alumnos. Uno de los aspectos más influyentes en el nivel de dificultad de un PAV es su estructura semántico-matemática (Carpenter y Moser, 1984; Greer, 1992; Vergnaud, 1991). Por ejemplo, los PAV de estructura aditiva (aquellos que pueden resolverse sumando o restando) pueden categorizarse como problemas de cambio, comparación, combinación e igualación (Carpenter y Moser, 1984; Heller y Greeno, 1978). En función del conjunto desconocido y de las relaciones (aditivas o sustractivas) existentes entre los conjuntos pueden establecerse las subcategorías descritas en el Gráfico I.

Algunos de estos problemas podrían resolverse de una manera rutinaria. Por ejemplo, el problema de Cambio 2 del Gráfico I puede resolverse utilizando la "estrategia de la palabra clave" (Hegarty et al., 1995), tomando los datos del problema, y la expresión "perdí" como clave para restar, sin comprender su estructura matemática. Sin embargo, los problemas difíciles solo pueden resolverse comprendiendo las relaciones entre las cantidades (Verschaffel et al., 2000). Por ejemplo, para resolver el problema de Comparación 5 del Gráfico I, es necesario comprender que, si tengo más dinero que tú, tú tendrás menos que yo, y por lo tanto hay que restar, aunque la expresión "más que" sugiera que hay que sumar.

Recibido: 14/03/2023 Aceptado: 09/02/2024

GRÁFICO I. Tipos de estructuras semántico-matemáticas aditivas



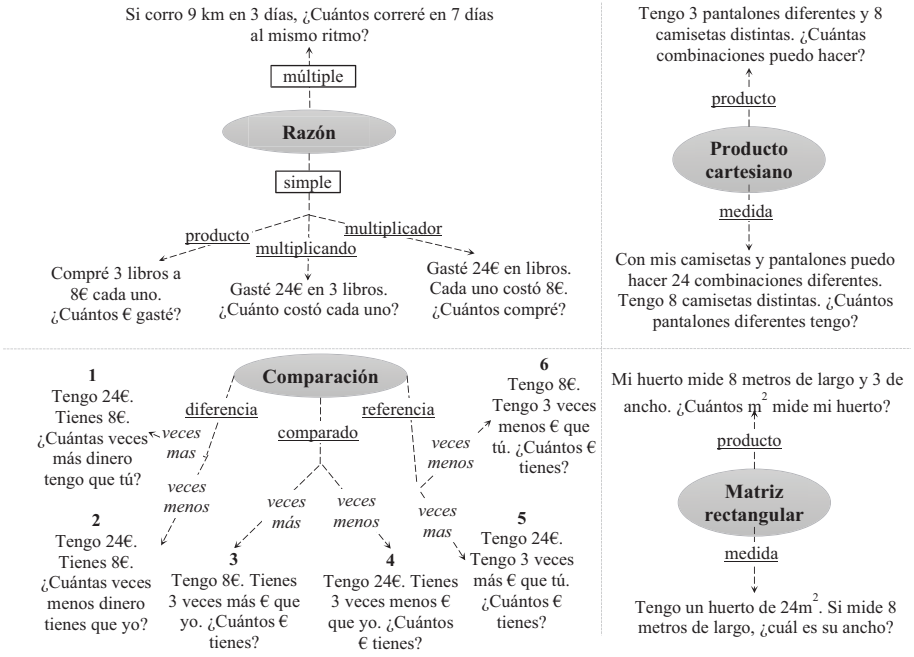
Fuente: Vicente et al. (2022). Elaboración propia.

De manera similar, en los PAV de estructura multiplicativa se han distinguido cuatro tipos (Greer, 1992; Vergnaud, 1991): razón (o grupos iguales), comparación multiplicativa (o escalares), producto cartesiano y matriz rectangular. Dependiendo de la operación requerida para resolver el PAV y del conjunto desconocido, se pueden establecer las subcategorías que se muestran en el Gráfico II.

Puesto que se han descrito 34 estructuras semántico-matemáticas para los PAV de estructuras aditivas o multiplicativas, y que la falta de práctica con algunos tipos de problemas parece dificultar su aprendizaje (Siegler y Oppenzato, 2021), los alumnos necesitan enfrentarse a la mayor variedad posible de problemas para aprender a resolverlos. De hecho, TIMSS incluye problemas de diferentes estructuras y niveles de dificultad

Recibido: 14/03/2023 Aceptado: 09/02/2024

GRÁFICO II. Tipos de estructuras semántico-matemáticas multiplicativas



Fuente: Vicente et al. (2022). Elaboración propia.

en sus evaluaciones (I.N.E.E., 2016). Siendo esto así, un análisis de los libros de texto puede aportar información relevante sobre la variedad de problemas a la que se enfrentan los alumnos en clase de matemáticas.

Libros de Texto y Resolución de Problemas

El libro de texto se considera un recurso didáctico muy influyente en el aprendizaje escolar por tres motivos. Primero, porque lo utilizan la gran mayoría de profesores (el 94% de los docentes de los países pertenecientes a la Organización para la Cooperación Económica y el Desarrollo, Mullis et al., 2008). En España, el 81,3% de los profesores reconocen emplearlo a diario (ANELE, 2014); otros estudios informan de un uso mayor en comunidades autónomas como Madrid (97.14%,

Fernández et al., 2013) o el País Vasco (96%, Mullis et al., 2008). En segundo lugar, porque traduce el currículo oficial a una secuencia de acciones concretas que profesores y alumnos pueden seguir, actuando como mediador entre el currículo oficial y el implementado (Valverde et al., 2002), y determinando en gran medida lo que se enseña en el aula (Oates, 2014). En tercer lugar, porque influyen en el aprendizaje de los alumnos; en relación con las matemáticas, los alumnos son más competentes en los contenidos a los que los libros dedican más espacio (Schmidt et al., 2001), más ejercicios y más problemas (Törnroos, 2005); y aprenden mejor los principios aritméticos y utilizan con más frecuencia las estrategias de resolución de problemas que más aparecen en los libros (Fagginger et al., 2016; Sievert et al., 2019, 2021). En cambio, los alumnos son poco competentes resolviendo determinados problemas de fracciones y decimales que, siendo matemáticamente simples, apenas aparecen en los libros (Siegler y Oppenzato, 2021).

Algunas variables específicas de los libros de texto pueden influir en el rendimiento de los alumnos, como determinados factores socio-económicos (p.e., el tipo de centro), el soporte (digital o papel) o el diseño de aspectos clave del libro, como la distribución de las tareas (Behnke, 2018). En relación con la resolución de problemas, algunos estudios han mostrado que los libros de países orientales como Japón, China o Singapur (cuyos alumnos han demostrado un elevado nivel de competencia matemática) contienen una distribución más diversificada y equilibrada de PAV aditivos y multiplicativos que los de países como Estados Unidos (Schoenfeld, 1991; Stigler et al., 1986; Xin, 2007) y España (Orrantia et al., 2005; Tárraga et al., 2021; Vicente et al., 2018), cuyos alumnos son menos competentes.

En resumen, los programas de evaluación de rendimiento matemático como TIMSS muestran las diferencias de rendimiento en matemáticas y resolución de problemas existentes entre los alumnos de diferentes contextos geográficos y culturales. Estas diferencias podrían deberse, en parte, a los libros de texto, puesto que los profesores los utilizan muy frecuentemente, y los alumnos aprenden mejor los contenidos que más aparecen en ellos. Por ello, para aprender a resolver PAV de diferentes niveles de dificultad los alumnos necesitarían libros que incluyeran una cantidad y una variedad amplia de problemas. De hecho, los libros de algunos países, cuyos alumnos han mostrado una mayor habilidad para resolver problemas, incluyen PAV más variados y de mayor dificultad que

Recibido: 14/03/2023 Aceptado: 09/02/2024

los libros de otros países, cuyos alumnos parecen ser menos competentes. De este modo, sería interesante analizar las diferencias existentes en los libros más utilizados en dos comunidades autónomas españolas, cuyos alumnos han mostrado niveles de competencia matemática muy dispares (CyL y Andalucía), en relación con los libros más utilizados en cada comunidad, y con la presencia y variedad de estructuras semántico-matemáticas de los PAV incluidos en esos libros.

El Presente Trabajo

Se pretende responder a la pregunta de si existen diferencias en los PAV incluidos en los libros más utilizados en las aulas de Primaria de CyL y Andalucía, mediante dos estudios diferentes. En el primer estudio se comprobará si los libros de texto más utilizados en las aulas de 3º de Educación Primaria[1] en CyL son diferentes a los más utilizados en Andalucía. En el segundo estudio se comprobará si existen diferencias en los PAV de los distintos libros de acuerdo con tres aspectos: a) porcentaje de actividades destinadas a resolver PAV frente a otras actividades matemáticas; b) variedad de estructuras semántico-matemáticas; y c) porcentaje de problemas según su nivel de dificultad semántico-matemática.

Diseño y metodología

Muestra

El universo del estudio 1 estaba compuesto por los libros de matemáticas utilizados en las clases de 3º de Primaria de los centros públicos y concertados de CyL (851) y Andalucía (2460), durante el curso 2020-2021. Se identificaron los libros de 709 centros de CyL (83.31% del total) y de 2425 de Andalucía (98.57%). Para el estudio 2 se seleccionaron

[1] Se analizaron los libros de un único curso para que la muestra fuera abarcable, considerando el número de variables y de libros analizados. Se eligió 3º de Primaria porque en ese nivel aparecen estructuras tanto aditivas como multiplicativas, y no son frecuentes los problemas con fracciones, decimales o transformaciones de unidades de medida, que implican niveles adicionales de dificultad.

Recibido: 14/03/2023 Aceptado: 09/02/2024

los libros que, siendo los más utilizados en cada comunidad autónoma, se emplearan en conjunto en, al menos, el 80% de los centros de cada comunidad.

Procedimiento

En el estudio 1 se determinaron los libros que se utilizaban en CyL consultando las páginas web de todos los centros de Educación Primaria de la comunidad, disponibles en la web de la Consejería de Educación. Aquellos centros que no publicaron el listado de libros en su web fueron contactados por correo electrónico. En el caso de Andalucía, los libros estaban publicados en la web de la Consejería de Educación. Después se calculó el porcentaje del total de centros en los que se utilizaba cada libro en cada comunidad autónoma. Para preservar la privacidad de las editoriales, se codificó el nombre de cada editorial con una combinación de dos números, el primero referido a la editorial, y el segundo al proyecto editorial. Posteriormente se realizó una segunda codificación de los libros más frecuentes en cada comunidad autónoma, que representara el orden que ocupa respecto a su porcentaje de uso. Para CyL se asignaron letras correlativas, y números en el caso de Andalucía. Los libros utilizados en ambas comunidades fueron codificados con una combinación de letra y número. Esta segunda codificación fue la utilizada para el estudio 2.

En relación con el estudio 2, primero se analizaron todas las actividades de los libros, entendiendo como "actividad" una tarea o conjunto de tareas agrupadas bajo un mismo epígrafe, en las que el alumno debía responder a una o varias preguntas que requerían cálculos o la aplicación de conocimientos matemáticos. Cada actividad se categorizó como: a) resolución de PAV; o b) otras actividades (en adelante OAs). Se categorizaron como actividades de resolución de PAV aquellas que incluían uno o varios PAV, considerándose PAV aquellas tareas en las que: a) se realizaba una descripción verbal de situaciones reales o imaginarias y se planteaba una pregunta matemática que se respondía aplicando al menos una de las cuatro operaciones aritméticas básicas; y b) podían clasificarse en alguna de las estructuras descritas en los Gráficos I y II. Se consideraron OAs el resto de actividades, incluyendo otros tipos de problemas verbales (p.e., estadísticos, geométricos) no clasificables en ninguna de las estructuras semántico-matemáticas de nuestro estudio. Se

Recibido: 14/03/2023 Aceptado: 09/02/2024

identificaron 8105 actividades, de las cuales 2057 fueron actividades de resolución de PAV, que incluyeron 3834 PAV.

En segundo lugar, se analizó la variedad de estructuras semántico-matemáticas. Se utilizaron clasificaciones diferentes para analizar las estructuras aditivas y multiplicativas; los primeros se clasificaron como problemas de cambio, comparación, combinación e igualación (Carpenter y Moser, 1984; Heller y Greeno, 1978), estableciéndose 20 subcategorías diferentes según el conjunto desconocido y las relaciones aditivas o sustractivas que se establecieran entre las cantidades del problema (ver Gráfico I). Los PAV de estructura multiplicativa se clasificaron como problemas de razón (o grupos iguales), comparación (o escalares), producto cartesiano y matriz rectangular (Greer, 1992; Vergnaud, 1991), estableciéndose 14 subcategorías diferentes (ver Gráfico II).

Finalmente, se categorizó cada estructura respecto al nivel de dificultad semántico-matemática, siguiendo el modelo de estrategias de resolución de Riley y Greeno (1988) para las estructuras aditivas, y los planteamientos de Greer (1992) y Vergnaud (1991) para las multiplicativas. Se consideraron fáciles las estructuras aditivas de cambio 1 y 2, comparación 1 y 2, igualación 1 y 2, y combinación 1; y las estructuras multiplicativas de razón simple, ya que son los más próximas a las de estructura aditiva (pueden llegar a resolverse como sumas o restas repetidas). Se consideraron de dificultad media las estructuras aditivas de cambio 3 y 4, comparación 3 y 4, igualación 5 y 6 y combinación 2 (ver Riley y Greeno, 1988, para las estructuras de combinación 2); y las multiplicativas de razón múltiple y las de comparación multiplicativa consistente, es decir, en las que los términos del problema 'veces más' o 'veces menos' coincidiera con la operación necesaria para resolverlo (una multiplicación y una división, respectivamente, ver Xin, 2007). Finalmente, se consideraron difíciles las estructuras aditivas de cambio 5 y 6, comparación 5 y 6, y de igualación 3 y 4, y las multiplicativas de comparación inconsistente, matriz rectangular y producto cartesiano.

Codificación y análisis de datos

Los autores codificaron conjuntamente 10 unidades didácticas de diferentes libros, para establecer los criterios de categorización. Después, dos de los autores analizaron otras 10 unidades didácticas,

tras lo cual se calculó el acuerdo inter-jueces utilizando el coeficiente de Kappa de Cohen, obteniéndose índices de correlación altamente significativos (κ =.98 para la clasificación de actividades de resolución de PAV y OAs; κ =.96 para las estructuras semántico-matemáticas de los PAV). Los dos primeros autores analizaron las estructuras de los PAV del resto de la muestra, y los tres últimos, categorizaron las actividades como PAV u OAs.

Respecto a las medidas utilizadas, para el estudio 1 se calculó el porcentaje de uso de cada editorial en cada comunidad autónoma. Para el estudio 2 se calcularon dos medidas. Primero, para cada variable, se calculó una medida ponderada de las frecuencias observadas en los libros según el porcentaje de uso en cada comunidad autónoma, mediante la siguiente fórmula:

$$M_{CA} = (F_1 \times P_1) + (F_2 \times P_2) + \ldots. (F_n \times P_n)$$

donde M_{CA} es la medida ponderada de los libros de una comunidad autónoma concreta, F es la frecuencia absoluta de cada libro en esa variable y P es el porcentaje de uso de ese libro en esa comunidad autónoma. Aunque esta medida implica una ligera sobreestimación de las frecuencias observadas en los libros andaluces (la suma de los porcentajes de uso es ligeramente superior en Andalucía que en CyL), esto no afecta a la validez de los resultados.

La segunda medida del estudio 2 fue la frecuencia observada en cada variable, en cada libro. Esta medida permitió realizar comparaciones entre todos los libros. Para facilitar la interpretación de los resultados, se calcularon los porcentajes correspondientes a cada frecuencia observada.

Para comprobar la significatividad estadística de las diferencias encontradas, en todos los análisis se utilizó el estadístico no paramétrico Chi-cuadrado para analizar las diferencias globales, a través de tablas planificadas, y el estadístico exacto de Fisher cuando el cálculo de Chi-cuadrado no era apropiado. Para verificar el efecto de esas diferencias se utilizó el estadístico V de Cramer el cual, de acuerdo con Cohen (1988), indica si el efecto de las diferencias es pequeño (.1), mediano (.3) o grande (.5). Para comparar las diferencias entre libros concretos se utilizaron pruebas z con un nivel de significatividad de .05, utilizando comparaciones por pares de proporciones de columna de las tablas planificadas.

Recibido: 14/03/2023 Aceptado: 09/02/2024

Hipótesis

En el estudio 1, de acuerdo con el marco teórico presentado, si los libros influyen en las diferencias de rendimiento en resolución de PAV entre los alumnos de CyL y Andalucía, los libros más utilizados en las clases de matemáticas de 3º de Primaria de CyL serán diferentes a los más utilizados en Andalucía (hipótesis 1).

En el Estudio 2, considerando los resultados de estudios previos sobre la influencia de los libros en el aprendizaje de las matemáticas, y las diferencias en las puntuaciones en TIMSS de los estudiantes castellano-leoneses y andaluces, cabría esperar que los libros más utilizados en CyL, comparados con los más utilizados en Andalucía, presentaran varias de las siguientes diferencias:

a) Un mayor porcentaje de actividades de resolución de PAV (hipótesis 2a).
b) Una mayor variedad de estructuras semántico-matemáticas (hipótesis 2b).
c) Un mayor porcentaje de PAV de estructuras de dificultad semántico-matemática media y alta, tanto aditivas (hipótesis 2c1) como multiplicativas (hipótesis 2c2).

Resultados

Estudio 1

En CyL se utilizaron 46 libros de 16 editoriales, mientras que en Andalucía se utilizaron 15 libros de 8 editoriales. Los libros más frecuentes en CyL fueron diferentes a los más utilizados en Andalucía, confirmándose la hipótesis 1. No se encontró una asociación significativa entre los libros utilizados en ambas comunidades autónomas ($p < .001$), siendo grande el tamaño de la diferencia (.81). En CyL los más utilizados fueron 1-1 (30.97%); 1-2 (12.78%); 2-1 (13.35%); y 3-1 (14.48%). En cambio, en Andalucía los libros más frecuentes fueron: 1-2 (26.18%); 1-3 (9.6%); 2-2 (31.92%); 3-2 (11.91%); y 4-2 (8.08%). Asimismo, el porcentaje de uso en CyL de 1-1, 2-1 y 3-1 fue mayor que en Andalucía; en cambio,

Recibido: 14/03/2023 Aceptado: 09/02/2024

en Andalucía se utilizaron con más frecuencia los libros 2 y 3 de las editoriales 1 y 2, y los libros 3-2, 4-2 y 6-1 (ver Tabla I).

Considerando los resultados del estudio 1, para el estudio 2 se seleccionaron los 5 libros más utilizados en cada comunidad autónoma. Esos libros se codificaron siguiendo el procedimiento anteriormente descrito (ver Tabla I). Cabe destacar que dos libros frecuentemente utilizados en Andalucía (4-2 y 1-3) también se utilizaron en CyL aunque no se encontraran entre las cinco editoriales más utilizadas en esa comunidad, por lo que también se incluyeron en el estudio 2. Dado que esos dos libros fueron la 10ª y la 7ª editoriales más utilizadas en CyL, respectivamente, fueron codificados con la 10ª y la 7ª letra ("J" y "G").

TABLA I. Frecuencias y porcentajes de uso de los libros más utilizados en 3º de Primaria en los centros de CyL y Andalucía

EDITORIAL	LIBRO	CyL		ANDALUCÍA		Código Estudio 2
		N	%	N	%	
1	**1-1**	218	30.97*	0	0	A
	1-2	90	12.78	635	26.19*	D2
	1-3	11	1.56	233	9.61*	J4
	1-4	7	1	0	0	
	Total	326	46.31	868	35.80	
2	**2-1**	94	13.35*	0	0	C
	2-2	31	4.40	774	31.92*	E1
	2-3	17	2.41	122	5.03*	
	Otros	4	0.56	0	0	
	Total	146	20.72	896	36.95	
3	**3-1**	102	14.49*	0	0	B
	3-2	34	4.83	289	11.92*	F3
	Otros	6	0.85	0	0	
	Total	142	20.17	289	11.92	
4	**4-1**	20	2.84	0	0	
	4-2	18	2.56	196	8.08*	G5
	Otros	6	0.85	0	0	
	Total	44	6.25	196	8.08	

(Continua)

Recibido: 14/03/2023 Aceptado: 09/02/2024

TABLA I. Frecuencias y porcentajes de uso de los libros más utilizados en 3º de Primaria en los centros de CyL y Andalucía *(Continuación)*

EDITORIAL	LIBRO	CyL		ANDALUCÍA		Código Estudio 2
		N	%	N	%	
5	**Sin especificar**	1	0.14	55	2.27*	
	Otros	6	0.86	0	0	
	Total	7	1	55	2.27	
6	**6-1**	1	0.14	118	4.87*	
	6-2	7	1	0	0	
	Total	8	1.13	118	4.87	
7	**7-1**	8	1.14	0	0	
OTRAS EDITORIALES		17	1.72*	3	0.11	
SIN LIBROS		11	1.56*	0	0	
TOTAL		709	100	2425	100	

Fuente: Elaboración propia.
Nota. En "Otras editoriales" se incluyeron los libros de aquellas editoriales con frecuencias absolutas menores que 5 o porcentajes inferiores al 0.7 del total de cada comunidad. El asterisco indica las diferencias estadísticamente significativas.

Estudio 2

Porcentaje de actividades de resolución de PAV

La frecuencia ponderada de actividades dedicadas a resolver PAV en los libros de CyL fue de 262, y de 693 a OAs (un 27.4% y un 72.6%, respectivamente). En los libros de Andalucía se observó una frecuencia ponderada de 207 tareas de resolución de PAV y de 654 OAs (un 24% y un 76%, respectivamente). Aunque la distribución de las tareas en los libros de ambas comunidades fue diferente -las actividades dedicadas a resolver PAV representaban un mayor porcentaje en los libros de CyL- esa diferencia no llegó a alcanzar la significatividad estadística.

No obstante, la distribución de tareas dedicadas a resolver PAV y OAs sí fue diferente en los distintos libros analizados, χ^2 (7, n = 8105) = 42.74, p <.01, aunque el efecto fue pequeño (.07). Los libros A y B (los más utilizados en CyL) dedicaron un mayor porcentaje de actividades

a resolver PAV; de hecho, esta diferencia fue significativa respecto al libro E1, el más utilizado en Andalucía (ver Tabla II). Asimismo, en los dos libros más frecuentes en CyL se encontraron frecuencias mayores de las esperadas, mientras que en el más frecuente de Andalucía se encontró una frecuencia menor de la esperada. Otros libros frecuentes en Andalucía, como J4 y G5, también incluyeron menos actividades de resolución de PAV que A, y frecuencias menores de las esperadas. Estos resultados confirman la hipótesis 2a. Finalmente, resulta llamativa la diferencia del total de tareas entre los diferentes libros, especialmente entre los libros A, y F3 y J4.

TABLA II. Frecuencias y porcentajes de actividades de resolución de PAV y OAs por editorial.

LIBRO	PAV			OAs			N TOTAL
	N	Res.Est.	%	N	Res.Est.	%	
A	383	7.85	29.6$^{E1.\ J4.G5.}$	911	5.64	70.4	1294
B	289	1.99	27.9^{E1}	748	-0.29	72.1	1037
C	269	0.74	24	850	3.42	76	1119
D2	307	3.11	28.3$^{E1.J4}$	776	0.73	71.7	1083
E1	216	-2.56	20.7	829	2.65	79.3$^{A.B.D2}$	1045
F3	195	-3.87	26	556	-7.27	74	751
J4	150	-6.68	21.6	546	-7.64	78.4$^{A.D2}$	696
G5	248	-0.57	23	832	2.76	77^{A}	1080

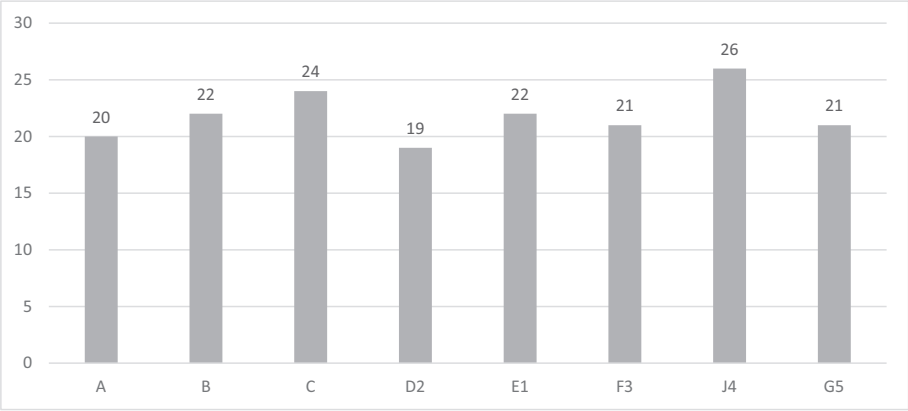
Fuente: Elaboración propia
Nota: Res.Est. = residuo estandarizado..

Variedad de estructuras semántico-matemáticas

La frecuencia ponderada de estructuras de los libros castellano-leoneses y andaluces fue muy similar (17.44 y 19.18, respectivamente); de hecho, la diferencia en la distribución de esas estructuras en los libros de una y otra comunidad autónoma no fue significativa. Tampoco se encontraron diferencias significativas en la distribución de las estructuras entre los libros, ni en las comparaciones realizadas entre pares de libros (ver Gráfico III), por lo que los resultados no confirman la hipótesis 2b.

Recibido: 14/03/2023 Aceptado: 09/02/2024

GRÁFICO III. Número de estructuras semántico-matemáticas incluidas por cada libro



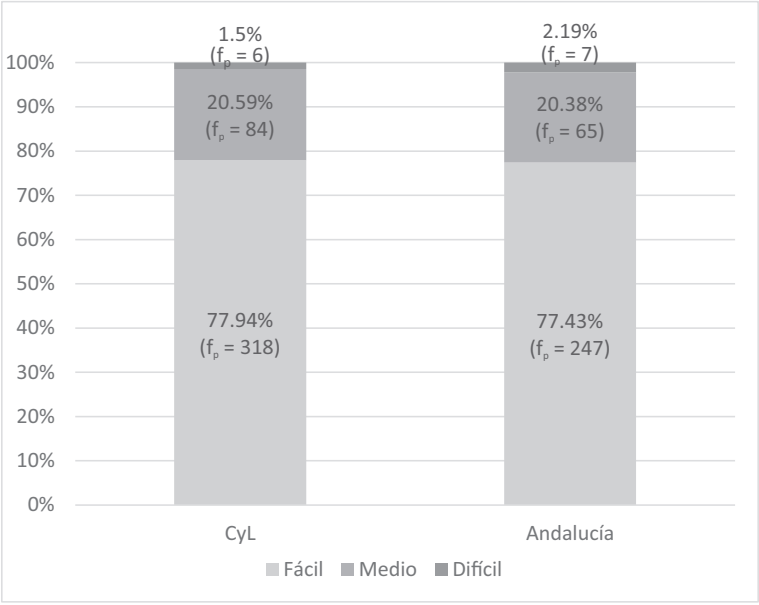
Fuente: elaboración propia

Nivel de dificultad semántico-matemática. Estructuras aditivas

La distribución de las frecuencias ponderadas de los tres niveles de dificultad fue muy similar en los libros de ambas comunidades, no encontrándose diferencias significativas. Los libros de CyL incluyeron más estructuras fáciles y de dificultad media, si bien los porcentajes que representaban respecto al total eran prácticamente idénticos (ver Gráfico IV).

La comparación entre los diferentes libros mostró que la gran mayoría de las estructuras aditivas fueron fáciles, mientras que las estructuras difíciles fueron muy escasas, salvo en el libro J4. La distribución de los niveles de dificultad en los distintos libros fue significativamente diferente entre ellos ($p < .001$), aunque el tamaño de esta diferencia fue pequeño (.16). No hubo diferencias significativas entre las editoriales más frecuentes en CyL y Andalucía (ver Tabla III). Las únicas diferencias significativas encontradas fueron las siguientes: a) los libros A, D2 y G5 incluyeron más problemas fáciles que B y J4, aunque solo en A y D2 se encontraron frecuencias mayores a las esperadas; b) el libro B incluyó más estructuras de dificultad media que A (aunque el residuo estandarizado de A fue mayor) y G5, donde se encontró una frecuencia muy por debajo de lo esperado (al igual que F3). Finalmente, respecto a las estructuras difíciles, el libro J4 incluyó

GRÁFICO IV. Frecuencias y porcentajes de estructuras aditivas por nivel de dificultad semántico-matemática, ponderadas por comunidad autónoma



Fuente: Elaboración propia.

TABLA III: Frecuencias y porcentajes de estructuras aditivas de cada nivel de dificultad semántico-matemática por libro

LIBRO	FACIL			MEDIO			DIFÍCIL			N TOTAL
	N	Res.Est.	%	N	Res.Est.	%	N	Res.Est.	%	
A	584	18.13	79.9^{B,J4}	141	7.77	19.3	6	-1.48	0.8	731
B	227	-3.19	68.8	92	2.08	27.9^{A,G5}	11	0.03	3.3	330
C	164	-6.95	77.4	40	-3.96	18.9	8	-0.88	3.8^{A}	212
D2	451	10.19	78.8^{B,J4}	117	4.98	20.5	4	-2.09	0.7	572
E1	226	-3.25	76.6	65	-1.06	22	4	-2.09	1.4	295
F3	132	-8.86	77.6	34	-4.66	20	4	-2.09	2.4	170
J4	248	-1.93	69.3	67	-0.82	18.7	43	9.72	12.0^{A,B,C,D2,E1,F3,G5}	358
G5	211	-4.14	82.7^{B,J4}	37	-4.31	14.5	7	-1.18	2.7	255

Fuente: elaboración propia.
Nota: Res.Est. = residuo estandarizado.

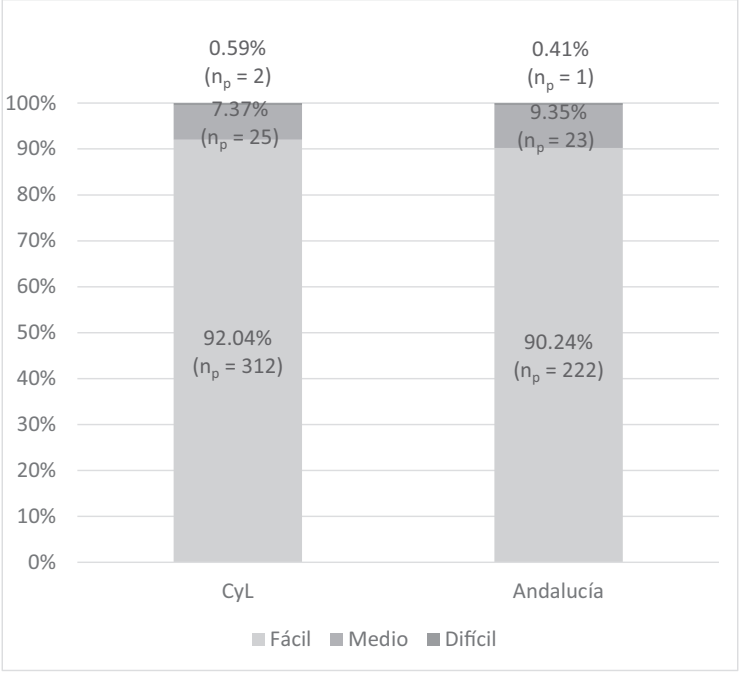
Recibido: 14/03/2023 Aceptado: 09/02/2024

más problemas difíciles que el resto, con una frecuencia muy por encima de lo esperado, mientras que el libro C incluyó más problemas difíciles que A. Estos resultados no confirman la hipótesis 2c1.

Nivel de dificultad semántico-matemática. Estructuras multiplicativas

De nuevo, las frecuencias ponderadas de los tres niveles de dificultad se distribuyeron de manera muy similar en los libros de ambas comunidades, sin que existieran diferencias significativas entre ambas distribuciones. Tal y como se muestra en el Gráfico V, en los libros de CyL se observó una frecuencia ponderada mayor de los tres niveles de dificultad que en los andaluces, aunque representaban prácticamente los mismos porcentajes respecto al total.

GRÁFICO V: Frecuencias y porcentajes de estructuras multiplicativas por nivel de dificultad semántico-matemática, ponderadas por comunidad autónoma



Fuente: Elaboración propia.

Respecto a la comparación entre los diferentes libros, todos mostraron una inmensa mayoría de estructuras fáciles y muy pocas estructuras difíciles. No obstante, la distribución de los niveles de dificultad fue diferente entre los libros analizados ($p < .001$), siendo pequeño el tamaño de esa diferencia (.14). No se encontraron diferencias significativas entre los libros más frecuentes en CyL y Andalucía (ver tabla IV). Se encontraron las siguientes diferencias: en primer lugar, los libros A y D2 –los dos con frecuencias muy por encima de lo esperado- incluyeron más estructuras fáciles que B, E1, F3 y J4; además, C y G5 incluyeron más estructuras fáciles que B y J4, siendo este último el que obtuvo una frecuencia más alejada de lo esperado. En segundo lugar, J4 incluyó más estructuras de dificultad media que el resto excepto B y E1, aunque su diferencia respecto al valor esperado fue menor que B; además, B y E1 también incluyeron más estructuras de dificultad media que A, C, D2 y G5, siendo esta última la que se quedó más lejos –por debajo- del valor esperado. Tomados en conjunto, estos resultados no confirman la hipótesis 2c2.

TABLA IV: Frecuencias y porcentajes de estructuras multiplicativas de cada nivel de dificultad semántico-matemática por libro.

LIBRO	FACIL			MEDIO			DIFÍCIL			N TOTAL
	N	Res.Est.	%	N	Res.Est.	%	N	Res.Est.	%	
A	522	15.47	94.2$^{B.E1.F3.J4}$	32	1.15	5.8	0	-1.55	0	554
B	267	-0.09	84.8	43	3.31	13.7$^{A.C.D2.G5}$	5	1.68	1.6	315
C	248	-1.25	93.6$^{B.J4}$	13	-2.56	4.9	4	1.03	1.5	265
D2	440	10.47	93.8$^{B.\ E1.F3.J4}$	29	0.57	6.2	0	-1.55	0	469
E1	186	-5.03	86.5	29	0.57	13.5$^{A.C.D2.G5}$	0	-1.55	0	215
F3	169	-6.07	86.2	23	-0.61	11.7	4	1.03	2	196
J4	113	-9.49	76.4	31	0.96	20.9$^{A.C.D2.G5}$	4	1.03	2.7	148
G5	203	-4.00	94.9$^{B.J4}$	9	-3.35	4.2	2	-0.26	0.9	214

Fuente: Elaboración propia.
Nota: Res.Est. = residuo estandarizado.

Recibido: 14/03/2023 Aceptado: 09/02/2024

Discusión

Según los resultados de TIMSS (I.N.E.E., 2016), los alumnos de CyL obtienen mejores puntuaciones en la prueba de matemáticas que los andaluces, por lo que, de acuerdo con los marcos teóricos del propio TIMSS, serían más hábiles para resolver PAV de diferentes niveles de dificultad (Mullis et al., 2020). Esta diferencia de rendimiento, que ha sido atribuida a factores generales como diferencias en el ISEC o un sistema educativo más equitativo (I.N.E.E., 2023), también podría deberse a los libros de texto. Los libros son el recurso didáctico más utilizado por los profesores para enseñar a resolver problemas a sus alumnos en el aula (Hiebert et al., 2003; Mullis et al., 2008). Además, la variedad y el nivel de dificultad de los problemas a los que se enfrentan los niños a partir de los libros influye en su habilidad para resolverlos (Siegler y Oppenzato, 2021), siendo la estructura semántico-matemática de esos problemas una de las variables que más influye en su nivel de dificultad (Carpenter y Moser, 1984; Greer, 1992; Vergnaud, 1991). Por otra parte, algunos estudios han mostrado que los libros de países con alumnos más competentes incluyen problemas de estructuras semántico-matemáticas variadas (Schoenfeld, 1991; Stigler et al., 1986; Xin, 2007), a diferencia de los libros de otros países, como España (Orrantia et al. 2005; Tárraga et al., 2021; Vicente et al., 2018). Siendo esto así, cabría preguntarse si parte de la diferencia en el nivel de rendimiento de los alumnos de CyL y Andalucía podría atribuirse a los libros de texto utilizados en una y otra comunidad autónoma. Si existiera tal influencia, cabría esperar que los libros más utilizados en CyL fueran diferentes a los más frecuentes en Andalucía y, en ese caso, que en los libros más frecuentes en CyL se encontraran: a) un mayor porcentaje de actividades de resolución de PAV, b) una mayor variedad de estructuras semántico-matemáticas, y c) un mayor porcentaje de PAV de dificultad semántico-matemática media o alta, de acuerdo con las estructuras y niveles de dificultad descritos en la literatura (Carpenter y Moser, 1984; Greer, 1992; Riley y Greeno, 1988; Vergnaud, 1991). Cuantas más diferencias de las anteriormente señaladas concurran, mayor influencia cabría atribuirle al diseño de los libros en las diferencias de rendimiento de los alumnos castellano-leoneses y andaluces.

Los resultados del estudio 1 han mostrado que en CyL y Andalucía se utilizan libros diferentes. No obstante, de las tres medidas consideradas

en el estudio 2 para analizar los PAV de esos libros, solo el porcentaje de actividades dedicadas a resolver PAV apunta a que los libros más utilizados en CyL son más adecuados que los de Andalucía para que los alumnos aprendan a resolver problemas. Si, además del porcentaje, tenemos en cuenta que la cantidad de actividades dedicadas a resolver PAV de los libros más utilizados en CyL es mucho mayor que en los de Andalucía (más del doble, en algunos casos), podría concluirse que esa práctica más intensa en resolución de problemas podría ser determinante. De estos resultados se desprende que podría no ser necesario un porcentaje elevado de problemas con dificultad media o alta (los libros más utilizados en CyL no incluyeron una variedad significativamente mayor de estructuras semántico-matemáticas, ni más problemas de dificultad semántico-matemática media o alta), pero sí mucha práctica con problemas fáciles y solo en algunos casos PAV más complejos. Estos resultados matizan los obtenidos en estudios previos que comparaban los problemas de los libros de diferentes países (Orrantia et al., 2005; Schoenfeld, 1991; Stigler et al., 1986; Tárraga et al., 2021; Xin, 2007), y coinciden en parte con los de Vicente et al. (2022) al comparar libros españoles con los de Singapur. Estos autores encontraron que los libros de Singapur contenían una mayor proporción de actividades de resolución de PAV que los españoles, y también problemas más difíciles, si bien el efecto de esta última diferencia era pequeño. En los libros singapurenses, al igual que en los analizados en nuestro estudio, la gran mayoría de los problemas eran fáciles. Las diferencias entre los libros de ambos países parecen residir en las ayudas al razonamiento que proporcionan a los alumnos, como el uso de representaciones gráficas (Kaur, 2015; Vicente, 2022), modelos de resolución más completos y basados en el razonamiento (Vicente et al., 2020) o un mayor nivel de autenticidad de los problemas (Vicente et al., 2021).

Asimismo, es razonable pensar que la influencia de los libros en el rendimiento de los alumnos esté mediada por el modo en el que los profesores resuelven los problemas en clase. Existen evidencias de que solo los profesores con un mayor conocimiento para la enseñanza de resolución de problemas promueven un mayor razonamiento al resolver cualquier tipo de problema; en cambio, los profesores con bajo conocimiento tienden a evitar los problemas más difíciles (Ramos et al., 2024). Siendo esto así, aunque algunos de los libros más utilizados en Andalucía contienen más problemas de dificultad media, quizá muchos no lleguen a implementarlos en el aula porque los consideren inadecuados para sus alumnos.

Recibido: 14/03/2023 Aceptado: 09/02/2024

Conclusiones

Los resultados del presente estudio permiten extraer tres conclusiones. Primero, aunque los libros de texto más utilizados en CyL no son los mismos que en Andalucía, todos esos libros son, en esencia, muy similares. De las tres medidas analizadas, solo el porcentaje de actividades de resolución de PAV parece marcar una diferencia clara entre los libros más utilizados en CyL y Andalucía. La ausencia de diferencias en la variedad de estructuras y niveles de dificultad semántico-matemática, tanto en los problemas aditivos como multiplicativos, llevan a pensar que la influencia de los libros en las diferencias de rendimiento de los alumnos de CyL y Andalucía parecen limitarse a una mayor cantidad de práctica con los PAV en los libros castellano-leoneses.

Una segunda conclusión es que, a pesar de que los libros más utilizados en CyL contienen PAV de un número de estructuras muy reducido y de baja dificultad semántico-matemática, parecen ser un material suficiente para que los alumnos desarrollen una habilidad para resolver problemas por encima de la media del país. Más allá de la cantidad de PAV que contienen, esto puede deberse, entre otras causas, a características de los libros no descritas en este estudio y al modo en que los maestros los utilizan en el aula.

Una tercera conclusión es que algunos libros podrían ser más adecuados que el resto para la enseñanza de determinados aspectos relacionados con los PAV, puesto que, aunque no se han encontrado grandes diferencias entre los libros más utilizados en CyL y Andalucía, sí existen diferencias entre libros concretos. Por ejemplo, mientras A o D2 proporcionan un banco amplio de problemas, y problemas de mayor nivel de complejidad procedimental, J4 presenta tanto una variedad mayor de estructuras como problemas más difíciles a nivel semántico-matemático (también B, para los problemas de estructuras aditivas).

Limitaciones y prospectiva

Los resultados del presente estudio deben interpretarse con cautela debido a la existencia de una serie de limitaciones. En primer lugar, la muestra de libros analizada es limitada, puesto que solo se incluyeron libros de 3º de Primaria de dos comunidades autónomas. Siendo esto así,

sería recomendable realizar estudios adicionales que amplíen la muestra de libros analizados a todos los cursos de Primaria, a otras comunidades autónomas e, incluso, a otros países. De hecho, sería conveniente replicar este estudio comparando los libros de diferentes regiones de otros países. En segundo lugar, de todas las fuentes de complejidad de los problemas solo se ha considerado la dificultad semántico-matemática, dejándose al margen otras variables de los problemas, como sus características lingüísticas, su carácter estereotipado o no estereotipado, o la ubicación de la incógnita. Estas cuestiones deberían analizarse en el futuro, junto con otros aspectos de los PAV como las ayudas a la resolución o su nivel de autenticidad. En tercer lugar, se han analizado los PAV incluidos en los libros, pero no su rol dentro de la unidad didáctica. Sería interesante conocer qué papel juega el pequeño porcentaje de problemas de dificultad semántico-matemática media o alta en las unidades didácticas (p.e., como desafío al principio de la unidad didáctica, como práctica al final de esta). Por último, sería muy interesante comparar el modo en el que profesores de ambas comunidades autónomas utilizan los libros en clase (p.e., su metodología de enseñanza o el modo de evaluar los aprendizajes).

Referencias bibliográficas

ANELE. (2014). *La edición de libros de texto en España*. ANELE.

Behnke, Y. (2018). Textbook Effects and Efficacy. En E. Fuchs & A. Bock (Eds), *The Palgrave Handbook of Textbook Studies*. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-53142-1_28

Carpenter, T.P., & Moser, J.M. (1984). The acquisition of addition and subtraction concepts. En R. Lesh y M. Landau (Eds.), *The acquisition of mathematical concepts & processes*, (pp. 7–44). Academic Press.

Cohen, J. (1988). *Statistical power and analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Inc. https://doi.org/10.1002/bs.3830330104

Fagginger, M., Hickendorff, M., van Putten, C. Beguin, A., & Heiser, W. (2016). Multilevel latent class analysis for large-scale educational assessment data. Exploring the relation between the curriculum and students' mathematical strategies. *Applied Measurement in Education, 29*(2), 144–159. https://doi.org/10.1080/08957347.2016.1138959

Recibido: 14/03/2023 Aceptado: 09/02/2024

Fernández, P., Caballero, P., & Fernández, J. (2013). ¿Yerra el niño o yerra el libro de Matemáticas? *Números. Revista de Didáctica de las Matemáticas, 83*, 131-148.

Greer, B. (1992). Multiplication and division as models of situations. En D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*, (pp. 276–295). Macmillan.

Hegarty, M., Mayer, R.E., & Monk, C.A. (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*, 18–32. https://doi.org/10.1037/0022-0663.87.1.18

Heller, J., & Greeno, J. (1978). *Semantic processing in arithmetic word problem solving*. Comunicación presentada en la Midwestern Psychological Association Convention, Chicago.

Hiebert, J., Gallimore, R., Givvin, K.B., Hollingsworth, H., Jacobs, J., Chui, A.M. et al. (2003). *Teaching mathematics in seven countries. Results from the TIMSS 1999 Video Study*. National Center for Education Statistics (NCES).

Hippe, R., Jakubowski, M., & Araújo, L. (2018). *Regional inequalities in PISA: the case of Italy and Spain*. Publications Office of the European Union. https://doi.org/doi:10.2760/495702

Instituto Nacional de Evaluación Educativa (2016). *TIMSS 2015. Estudio internacional de tendencias en Matemáticas y Ciencias. (Informe español: resultados y contexto)*. Ministerio de Educación, Cultura y Deporte. https://sede.educacion.gob.es/publiventa/descarga.action?f_codigo_agc=18230

Instituto Nacional de Evaluación Educativa (2023). *PISA 2022. Programa para la evaluación Internacional de los estudiantes. Informe español*. Ministerio de Educación, Formación Profesional y Deportes. https://www.libreria.educacion.gob.es/ebook/184935/free_download/

Kaur, B. (2015). The model method: A tool for representing and visualizing relationships. En X. Sun, B. Kaur & J. Novotna (Eds.), *Conference proceedings of ICMI Study 23: Primary mathematics study on whole numbers,* (pp. 448-455). http://www.umac.mo/fed/ICMI23/doc/ProceedingsICMI_STUDY_23_final.pdf.

Lindquist, M., Philpot, R., Mullis, I., & Cotter, K. E. (2017). TIMSS 2019 Mathematics Framework. En I.V.S. Mullis & M.O. Martin (Eds), *TIMSS 2019 Assessment Frameworks*. Recuperado de: http://timssandpirls.bc.edu/timss2019/frameworks/

Mullis, I., Martin, M., & Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grade*. TIMSS and PIRLS International Study Center, Boston College. http://pirls.bc.edu/timss2007/mathreport.html

Mullis, I., Martin, M., Foy, P., Kelly, D., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. TIMSS and PIRLS International Study Center, Boston College. https://timssandpirls.bc.edu/timss2019/international-results/

Oates, T. (2014). *Why textbooks count*. Cambridge assessments. http://www.cambridgeassessment.org.uk/Images/181744-why-textbooks-count-tim-oates.pdf.

Orrantia, J., González, L.B., & Vicente, S. (2005). Analysing arithmetic word problems in Primary Education textbooks. *Journal for the Study of Education and Development, 28*(4), 429–451. https://doi.org/10.1174/021037005774518929

Philpot, R., Lindquist, M., Mullis. I.V.S. & Aldrich, Ch. E.A. (2021). TIMSS 2023 Mathematics Framework. En I.V.S. Mullis, M.O. Martin & von Davier, M. (Eds), *TIMSS 2023 Assessment frameworks*. TIMSS and PIRLS International Study Center.

Ramos, M., Vicente, S., Rosales, J., & Chamoso, J. (2024). Influencia del conocimiento pedagógico del profesorado en su práctica de aula cuando resuelve problemas aritméticos verbales con su alumnado. Un estudio exploratorio. *Journal for the Study of Education and Development, 47*(2), 321-345. https://doi.org/10.1177/02103702241253415.

Rao, N., Ng, S. S. N., & Pearson, E. (2010). Preschool pedagogy: A fusion of traditional Chinese beliefs and contemporary notions of appropriate practice. En C. Chan, & N. Rao (Eds.), *Revisiting the Chinese learner. CERC studies in comparative education*, (pp. 255–279). Springer. https://doi.org/10.1007/978-90-481-3840-1_9

Riley, M., & Greeno, J. (1988). Developmental analysis of understanding language about quantities of solving problems. *Cognition and Instruction, 5*, 49–101. https://doi.org/10.1207/s1532690xci0501_2

Schmidt, W., McKnight, C., Houang, R., Wang, H., Wiley, D., Cogan, L., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. Bass.

Schoenfeld, A.H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics.

Recibido: 14/03/2023 Aceptado: 09/02/2024

En J.F. Voss, D.N. Perkins & J.W. Segal (Eds.), *Informal reasoning and education*, (pp. 311–343). Lawrence Erlbaum Associates.

Siegler, R., & Oppenzato, C. (2021). Missing Input: How Imbalanced Distributions of Textbook Problems Affect Mathematics Learning. *Child Development Perspectives, 15*(2), 76-82. https://doi.org/10.1111/cdep.12402

Sievert, H., van den Ham, A.K., & Heinze, A. (2021). Are first graders' arithmetic skills related to the quality of mathematics textbooks? A study on students' use of arithmetic principles. *Learning and Instruction, 71*(101401), 1–14. https://doi.org/10.1016/j.learninstruc.2020.101401

Sievert, H., van den Ham, A.K., Niedermeyer, I., & Heinze, A. (2019). Effects of mathematics textbooks on the development of primary school children's adaptive expertise in arithmetic. *Learning and Individual Differences, 74*(101716), 1–13.

Stigler, J., Fuson, K., Ham, M., & Kim, M. (1986). An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks. *Cognition and Instruction, 3*, 153–171.

Tárraga, R., Tarín, J., & Lacruz, I. (2021). Analysis of Word Problems in Primary Education Mathematics Textbooks in Spain. *Mathematics, 9*(17), 2123. https://doi.org/10.3390/math9172123

Törnroos, J. (2005). Mathematics textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation, 31*(4), 315–327.

Valverde, G., Bianchi, L. J., Wolfe, R., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks.* Kluwer Academic Publishers.

Vergnaud, G. (1991). *El niño, las matemáticas y la realidad.* Trillas.

Verschaffel, L., Depaepe, F., & Van Dooren, W. (2020). Word problems in mathematics education. En S. Lerman (Ed.), *Encyclopedia of mathematics education*, (pp. 908–911). Springer.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems.* Swets y Zeitlinger Publishers. https://doi.org/10.1023/A:1004190927303

Vicente, S., Manchado, E., & Verschaffel, L. (2018). Solving arithmetic word problems. An analysis of Spanish textbooks. *Culture and Educación, 30*, 71-104. https://doi.org/10.1080/11356405.2017.1421606

Vicente, S., Sánchez, R., & Verschaffel, L. (2020). Word problem solving approaches in mathematics textbooks: a comparison between

Singapore and Spain. *European Journal of Psychology of Education, 35*, 567–587. https://doi.org/10.1007/s10212-019-00447-3

Vicente, S., Verschaffel, L., & Múñez, D. (2021). Comparación del nivel de autenticidad de los problemas aritméticos verbales de los libros de texto españoles y singapurenses. *Cultura y Educación, 33*(1), 106-133. https://doi.org/10.1080/11356405.2020.1859738

Vicente, S., Verschaffel, L, Sánchez, R., & Múñez, D. (2022). Arithmetic word problem solving. Analysis of Singaporean and Spanish textbooks. *Educational Studies in Mathematics, 111*, 375–397. https://doi.org/10.1007/s10649-022-10169-x

Xin, Y.P. (2007). Word problem solving tasks in textbooks and their relation to student performance. *The Journal of Educational Research, 6,* 347–359. https://doi.org/10.3200/JOER.100.6.347-360

Información de contacto: Santiago Vicente Martín. Universidad de Salamanca. Facultad de Educación. Dpto. Psicología Evolutiva y de la Educación. Paseo Canalejas 169. 37008, Salamanca. E-mail: sanvicente@usal.es

Recibido: 14/03/2023 Aceptado: 09/02/2024

La Formación Profesional Dual en los centros educativos andaluces desde la perspectiva del profesorado

Dual Vocational Training in Andalusian educational institutions from the teachers' perspective

https://doi.org/10.4438/1988-592X-RE-2024-406-642

María Teresa Pozo Llorente
https://orcid.org/0000-0002-8120-1474
Universidad de Granada

María de Fátima Poza Vilches
https://orcid.org/0000-0001-6186-9306
Universidad de Granada.

José Manuel Aguayo Moral
https://orcid.org/0000-0002-7304-961X
Universidad de Granada.

Resumen

A pesar de la creciente presencia de la Formación Profesional Dual (FPD) en el sistema educativo español y de los esfuerzos realizados desde distintos organismos por impulsarla, son muchas las voces que alertan de que el proceso de puesta en marcha de esta modalidad formativa no se está haciendo como debiera, lo que dificulta su éxito. El apoyo de las administraciones, su financiación y normativa, la pertinencia de la oferta, el papel de las empresas, el reconocimiento al profesorado o la coordinación centro-empresa son algunos de los aspectos que están detrás de las controversias que desde el principio han acompañado a esta modalidad formativa. El propósito de este estudio es caracterizar este debate en el escenario educativo andaluz a partir de la opinión de sus protagonistas diarios, el profesorado implicado, y de las prácticas que

Recibido: 22/03/2023 Aceptado:22/12/2023

se desarrollan en los centros andaluces. Se explora en qué medida el género, la experiencia en formación dual y la participación en gestión pueden estar condicionando esta opinión. Los docentes encuestados han sido 1.445, de los 2.164 implicados en FPD en el curso 2019-20. El cuestionario utilizado ha sido diseñado ad hoc y la opinión recogida se ha estructurado en torno a cinco dimensiones relacionadas con los tópicos de los debates nacionales. Los análisis realizados destacan que las prácticas desarrolladas en los centros andaluces relacionadas con la gestión administrativa de la dual y con el profesorado y la acción tutorial son las que más favorecen la implantación de la FPD en Andalucía y las relacionadas con el modelo de gobernanza, la organización académica y el papel de la empresa son las que presentan margen de mejora. Las profesoras, el profesorado implicado en gestión y el que tiene gran experiencia en el sistema dual son los que mejor valoran este sistema.

Palabras clave: Formación Profesional Dual, profesorado, estudio evaluativo, buenas prácticas, aspectos de mejora.

Abstract

Despite the increasing presence of Dual Vocational Education and Training in the Spanish education system and the great efforts made by different bodies to promote this training pathway, many concerns have been raised that perhaps its launch is not being conducted correctly, which may hinder its success. The support of the different levels of government, funding and regulations, the pertinence of the training offered, the role of the companies, the recognition of teachers, and company-school coordination are some of the issues around which controversies have arisen since this training model was first implemented. The purpose of this study is to characterize how this debate is playing out in the Andalusian education system, based on the opinion of its day-to-day protagonists –the teachers involved– and on the practices taking place in Andalusian schools. The paper explores to what extent the teachers' gender, their experience in dual training and their degree of involvement in administrative coordination activities may be conditioning these opinions. A total of 1,445 teachers took part in the study, out of the 2,164 teachers involved in dual training programmes in the 2019-20 academic year. The questionnaire was designed *ad hoc* and the questions were structured around five dimensions related to issues present in national debates. The analyses performed indicate that the practices most contribute to the successful implementation of this training model in Andalusian educational institutions are those related to the administrative coordination of dual training programmes and those related to teachers and guidance services. The practices related to the governance model, academic organisation and the role of the company are those that show the most room for improvement. Female teachers, teaching staff who are involved in management and teaching staff with greater experience in the dual system are those that rate the system most highly.

Recibido: 22/03/2023 Aceptado:22/12/2023

Keywords: Dual Vocational Education and Training, teachers, evaluative research, good practices, areas to improve.

Introducción

En los últimos años, en España, la Formación Profesional (en adelante FP) tanto del ámbito educativo como laboral, se ha convertido en uno de los ejes estratégicos de las políticas del gobierno y, su reforma, en una gran prioridad del país; una reforma cuyo objetivo es adaptar la FP a los cambios y demandas de la nueva economía y del tejido empresarial, garantizando así la fortaleza competitiva y la mejora de la empleabilidad en España. Muestra de esta apuesta gubernamental por la FP son los dos planes nacionales aprobados (MEFP, 2020a y 2020b) y la Ley Orgánica 3/2022, de 31 de marzo, de ordenación e integración de la FP (B.O.E., 2022) con los que se pretende su mejora y europeización.

En este reconocimiento social y empresarial ha tenido mucho que ver tanto la implantación, en el 2012, de la formación en alternancia entre el centro educativo y la empresa, como la prospección que en el 2020 realizó el Centro Europeo para el Desarrollo de la Formación Profesional (CEDEFOP, 2020), en la que destaca que entre los años 2018 y 2030 el 65% de los puestos de trabajo que se generen en España demandarán profesionales con cualificaciones medias (FP de Grado Medio) y el 35% restantes con cualificaciones altas (FP de Grado Superior y formación universitaria), además, se requerirán más perfiles de tipo técnico y con formación STEAM (Ciencia, Tecnología, Ingeniería, Arte y Matemáticas). Sin embargo, y como plantean Echeverría y Martínez-Clares (2021) aún queda camino por recorrer; como muestra el informe de la OCDE (MEFP, 2020c), España tiene la tasa más alta de abandono escolar temprano de la UE (17,9%), cuya media es del 10,6%.

Cuando el gobierno español apuesta por la modalidad de FP en alternancia lo hace desde el convencimiento de que esta modalidad de FP constituye una herramienta muy eficaz frente a algunos de los grandes problemas del sistema de FP. La FP Dual (en adelante, FPD) se incorpora con el propósito de aumentar el número de jóvenes que alcanza un título de educación postobligatoria, mejorar sus oportunidades de inserción profesional, disminuir la tasa de paro juvenil (en torno al 50% en el 2012) y adecuar su formación a los requerimientos de un contexto socioeconómico y laboral cada vez más exigente con los perfiles de

sus futuros profesionales y caracterizado por la implantación de pautas de comportamiento marcadamente tecnológicas determinadas por el contexto de la Industria 4.0. (Barrientos, et al., 2019; Igartua, 2017; Ishar, Derahman y Kamin, 2020; Vaganova, et al., 2020; entre otros). El Real Decreto 1529/2012 (B.O.E., 2012) se convierte en el marco regulador tanto del contrato para la formación y el aprendizaje, de gran tradición en España, como el de la formación dual dentro del sistema educativo español, en concreto, dentro de la Formación Profesional Inicial.

A nivel nacional, la FPD, entendida como "un programa formal de educación y formación en el cual el aprendizaje se realiza de manera alternada en el lugar de trabajo y en una institución educativa, obteniendo una cualificación reconocida" (Pineda, et al., 2018, p.35), está cada vez más presente, tal y como muestra el aumento de proyectos, centros implicados, empresas colaboradoras y alumnado matriculado.

A pesar de la larga trayectoria y gran reconocimiento internacional de esta modalidad formativa (Emmenegger y Seitzl, 2020; Stephen, 2020; Vaganova, et al., 2020) y de los esfuerzos realizados desde distintos organismos nacionales e internacionales por impulsarla, son muchas las voces que alertan de que el proceso de puesta en marcha en España no se está haciendo como debiera, lo que dificulta su éxito. Aunque su incorporación al sistema educativo no se planteó como una respuesta puntual a una situación de crisis, sino como una propuesta en firme y con carácter estructural, desde su puesta en marcha, hace más de una década, el sistema dual ha estado rodeado de controversias y debates que, como plantea Sanz de Miguel (2017) evidencian el amplio margen de mejora que este sistema tiene especialmente en el ámbito de la gobernanza, de la misión y visión.

El apoyo real y efectivo de las administraciones, su financiación, la concreción de la normativa reguladora, la adecuación de la oferta formativa al sistema productivo y a la oportunidad del tejido empresarial territorial, el papel desempeñado por las empresas, el reconocimiento al profesorado implicado, la coordinación centro-empresa, la capacidad de esta modalidad para abarcar todo el currículum oficial, la suficiencia de recursos o la idoneidad y heterogeneidad de los modelos de gobernanza entre las distintas comunidades autónomas son algunos de los aspectos destacados desde distintos ámbitos (científico, técnico, educativo, empresarial, sindical, etc.) como condicionantes de la implantación y éxito de este modelo en el contexto español (Aguilar, 2015; Bentolila y Jansen, 2019; Echeverría, 2016a y 2016b; Echeverría y Martínez-Clares,

Recibido: 22/03/2023 Aceptado:22/12/2023

2018; Igartua, 2017; Marhuenda, et al., 2017; Martín-Artiles et al., 2020; Martínez-Morales y Marhuenda, 2020; Navarro, 2019; Pineda et al., 2019; Šćepanović y Martín-Artiles, 2020; Vila y Chisvert, 2018; Virgós y Burguera, 2020; Virgós, Burguera y Pérez, 2022; entre otros).

Desde su puesta en marcha experimental en Andalucía en el curso 2013-14 hasta la actualidad, la FPD está cada vez más presente en el escenario político, educativo y empresarial andaluz como medida para paliar la situación de paro juvenil; como señala la Comisión Europea (2020) la crisis económica del 2007 contribuyó a que el 55,5% de la juventud española entre 15 y 24 años, el 66% en la región andaluza, se encontrase en situación de desempleo en 2013.

La FPD del sistema educativo en Andalucía, desde su inicio, se estructura en proyectos formativos que los centros educativos diseñan, conjuntamente con las empresas colaboradoras, y presentan para su aprobación o renovación a la convocatoria que anualmente pone en marcha el gobierno andaluz tomando como base lo establecido en el RD1529/2012 (B.O.E., 2012). Estos proyectos se desarrollan dentro de los Ciclos formativos de FP de nivel básico, medio y superior, ofertando plazas duales en ellos.

La confianza de los sucesivos gobiernos de la comunidad autónoma andaluza en este sistema en alternancia ha sido muy alta. Desde su puesta en marcha en el curso 2013-14, la formación dual en Andalucía ha crecido anualmente pasando de 11 proyectos que cubrían 6 de las 26 familias profesionales contempladas en el Catálogo Nacional de Cualificaciones Profesionales, 227 alumnos/as, 11 centros y 56 convenios con empresas en el curso 2013/14 a 772 proyectos, 14.559 alumnos/as, 332 centros y 11.806 convenios con empresas en el curso 2022/23. También es necesario destacar el aumento en el número de familias profesionales con plazas duales en Andalucía; en el curso 2022/23 se cubren 24 de las 26 familias profesionales, siendo las relacionadas con Vidrio y cerámica y Artes y artesanías las que todavía no están representadas en la oferta dual en Andalucía.

La investigación que se recoge en este artículo se ubica en el marco presentado anteriormente y forma parte de un estudio más amplio cuyo propósito último ha sido aportar argumentos que justifiquen la conveniencia de seguir apostando por la FPD en Andalucía y avanzar en la definición de unas directrices que orienten la planificación estratégica y sostenible del sistema dual en esta región aportando evidencias que faciliten la toma de decisiones para su mejora.

El objetivo del trabajo que se presenta es doble; por un lado, caracterizar, en opinión del colectivo docente, como facilitadoras o no del éxito de esta modalidad de FP, algunas de las prácticas que se están desarrollando en los centros educativos andaluces y, por otro, explorar en qué medida esta opinión está condicionada por variables tales como el género, la experiencia en el sistema dual y/o la participación en gestión del profesorado implicado en el mismo.

Método

La respuesta a este objetivo se ha abordado desde un planteamiento metodológico evaluativo; la información se ha recogido a través de un cuestionario *ad hoc* dirigido a todo el profesorado de Andalucía implicado en la FPD durante el curso 2019-20. Para ello fue necesaria la colaboración de la Secretaría General de Educación y Formación Profesional de la Consejería de Educación y Deporte (actual Consejería de Desarrollo Educativo y Formación Profesional) de la Junta de Andalucía, que a través de los/as coordinadores/as duales provinciales y desde su plataforma de gestión interna (Séneca), contactó con el profesorado implicado facilitándole el acceso al cuestionario, animando a la participación y asegurando el anonimato en sus respuestas. El instrumento utilizado incorpora una breve presentación del equipo, del propósito de este, del compromiso de confidencialidad y del consentimiento informado.

Muestra

Según los datos facilitados por dicha Consejería, el número de docentes implicados en FPD en el curso 2019-20 ascendía a 2.164, distribuidos tal y como se recoge en la Tabla I[1]. Considerando como exigencias de muestreo y representatividad un nivel de confianza del 95% y un error del 5%, era necesario contar con una muestra de 387 docentes distribuidos por provincias según consta en la fila "Muestra" de la misma tabla. En todas las provincias se ha superado el tamaño mínimo necesario para

[1]. El desajuste entre respuestas y profesorado identificado en algunas provincias es motivado por el desfase entre la solicitud de datos y la recogida de información.

Recibido: 22/03/2023 Aceptado:22/12/2023

TABLA I. Participación del profesorado por provincias

	Almería	**Cádiz**	**Córdoba**	**Granada**	**Huelva**	**Jaén**	**Málaga**	**Sevilla**
Profesores/as	121	360	292	263	116	153	338	521
Muestra	22	64	52	47	21	28	60	93
Nº respuestas	125	216	171	186	105	173	181	286
Porcentaje	103,31%	60,00%	58,90%	70,72%	90,52%	113,07%	53,55%	55,09%

Fuente: elaboración propia.

tener una muestra representativa; si bien, el porcentaje de participación es dispar, oscilando entre el 53,55% de la provincia de Málaga y el 100% de las provincias de Jaén y de Almería.

El número de docentes encuestados ha sido de 1.445; el 56,5% hombres (N=815) y el 43% (N=620) mujeres (el 0,5% de las personas encuestadas no ha referenciado el género). La mayoría está en el rango de edad 36-55 y su situación contractual es indefinida.

El 70% del profesorado encuestado pertenece a centros públicos, el 27% a centros con estas enseñanzas concertadas y el 3% a centros privados.

El profesorado encuestado imparte docencia en ciclos formativos pertenecientes a las 23 familias profesionales[2] presentes en Andalucía en el curso 2019-20. La mayoría lo hacen en ciclos de las familias de Comercio y marketing seguidos de los de las familias de Administración y gestión, Servicios socioculturales y a la comunidad y Electricidad y electrónica.

Servicios socioculturales y a la comunidad (19H/68M), Sanidad (17H/43M) e Imagen personal (4H/10M) son las familias con más profesoras que profesores. La mayor presencia de profesores se da en familias como Instalación y mantenimiento (24H/4M), Electricidad y electrónica (66H/9M), Actividades físicas y deportivas (12H/4M) y Fabricación mecánica (30H/5M).

Son tres los perfiles profesionales del profesorado encuestado:

- Gestión [39,93%]: Profesorado con experiencia en gestión, bien como miembro del equipo directivo de un centro o en la coordinación de proyectos duales.

[2] https://sepe.es/HomeSepe/Personas/formacion/certificados-profesionalidad/familias-profesionales.html

Recibido: 22/03/2023 Aceptado:22/12/2023

- Docencia [39,80%]: Profesorado encuestado que imparte docencia en ciclos con proyectos duales en el momento de la recogida de información y sin experiencia en gestión.
- Prospección [20,27%]: Profesorado que ha desempeñado tareas de prospección según las Instrucciones del 10 de octubre de 2019 de la Consejería de Educación y Deporte.

De especial interés para este estudio ha sido el índice de desarrollo profesional del profesorado encuestado; esta variable, diseñada en el marco de este estudio, hace referencia a su formación y experiencia en FPD. Este índice se ha calculado como suma de los valores que se muestran en la Tabla II.

TABLA II. Definición del índice de desarrollo profesional del profesorado encuestado

Variables	Observación	Valor
Años de experiencia docente en dual	1-2	0
	3-4	1
	5-6	2
Participación en búsqueda de empresas	No	0
	Sí	1
Número de proyectos	0	0
	1	1
	Más de 1	2
Estancias formativas	No	0
	Sí	1
Participación en formación permanente	No	0
	Sí	1
Conocimiento sobre FP DUAL	1 – muy deficiente	0
	2	0
	3	0
	4	1
	5 – muy bueno	3

Fuente: elaboración propia

Recibido: 22/03/2023 Aceptado:22/12/2023

Este índice solo se ha calculado para el profesorado que ha contestado a todos los ítems del cuestionario que integran su cálculo (N=806). Así pues, este profesorado se clasifica en tres grupos de desarrollo profesional: índice bajo[3] el 32,1% del profesorado (N=259); medio[4], el 57,9% (N=467) y alto[5], el 9,9% (N=80).

Instrumento

El cuestionario utilizado para la recogida de información ha sido diseñado por el equipo responsable del estudio. La información recogida en la primera parte del instrumento ha permitido caracterizar tanto al profesorado encuestado como el escenario educativo dual en el que este desarrolla su tarea docente. La segunda parte, integrada por cinco bloques (el último solo para el profesorado que ha desarrollado tareas de prospección), con opciones de respuesta escalar, recoge la opinión del profesorado encuestado sobre las prácticas que se desarrollan en los centros andaluces en los que se imparten ciclos duales.

En este artículo, se aborda el análisis del Bloque I del cuestionario (25 ítems) centrado en los aspectos que obstaculizan y favorecen la implantación de esta modalidad de FP en los centros andaluces. Estos aspectos objeto de valoración giran en torno a la normativa, al apoyo de la Administración, al diseño y la oferta, a la gestión administrativa y educativa y al papel de los agentes implicados.

La calidad del instrumento se ha asegurado a través de su fiabilidad y validez (de contenido y constructo). La validez del contenido se ha apoyado en la revisión de documentación normativa y científica sobre FPD, en el análisis de los instrumentos utilizados en estudios previos tanto del equipo investigador (Pozo-Llorente y Poza-Vilches, 2020) como de otros equipos y en la validación interjueces. En esta validación participaron 12 personas (expertos/as en el tópico, equipos directivos de centros educativos, profesorado y técnicos y políticos de la Consejería de Educación con competencias en FPD). Tras analizar la opinión de las personas consultadas sobre la pertinencia, claridad y relevancia del cuestionario piloto, se decide eliminar 4 ítems, modificar la distribución de

[3] Índice Bajo = valor entre 0 y 3
[4] Índice Medio = valores entre 4 y 7
[5] Índice Alto = valores entre 8 y 10

algunos en las dimensiones teóricas que integran cada bloque, quedando finalmente integrado por 84 ítems distribuidos en cinco bloques. En el Anexo I se muestran las dimensiones teóricas e ítems asociados que derivan de esta consulta.

La fiabilidad por consistencia interna del cuestionario, según el Coeficiente Alfa de Cronbach es elevada en todos los bloques del cuestionario (todos cercanos a 1), siendo para el Bloque I (25 items), objeto de este artículo, el valor de α=0,93.

Procedimiento

El cuestionario fue diseñado inicialmente para su aplicación presencial, pero las especiales circunstancias derivadas de la crisis sanitaria por la COVID19 dificultaron todo el proceso. Esta situación pandémica obligó a adaptar el cuestionario para su aplicación online y retrasó la recogida de información a los meses de junio y julio de 2020, evitando coincidir con el obligado proceso de adaptación de los centros escolares andaluces a la docencia online.

Se ha llevado a cabo un Análisis Factorial Exploratorio basado en el método de Componentes Principales rotación Varimax con Kaiser para extraer los factores dentro del Bloque I (25 ítems) objeto de este artículo. En las Tablas III, IV y V se muestra información referente a la KMO y Bartlett, la cual es muy alta ya que su valor es cercano a 1 así como la varianza explicada con los 5 componentes que articulan este análisis factorial y la matriz de componentes rotados.

TABLA III. Prueba de KMO y Bartlett

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		**,92**
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	7796,56
	gl	300
	Sig.	,00
	5 – muy bueno	3

Fuente: elaboración propia.

Recibido: 22/03/2023 Aceptado:22/12/2023

TABLA IV. Varianza total explicada

Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	9,63	38,53	38,53	9,63	38,53	38,53	3,58	14,32	14,32
2	1,76	7,03	45,57	1,76	7,03	45,57	3,50	13,99	28,31
3	1,46	5,86	51,43	1,46	5,86	51,43	3,07	12,30	40,61
4	1,13	4,54	55,97	1,13	4,54	55,97	2,74	10,97	51,58
5	1,04	4,16	60,12	1,04	4,16	60,12	2,14	8,55	60,12
Método de extracción: análisis de componentes principales.									

Fuente: elaboración propia.

Tras la interpretación de los factores observados y en base a los ítems que conforman cada uno de ellos se redefinieron los agrupamientos inicialmente definidos (dimensiones) dentro de este bloque.

A partir de este análisis se ha procedido a realizar un Análisis Factorial Confirmatorio detallado en la Figura I y Tabla VI.

Este análisis confirma una correlación fuerte entre dimensiones (Tabla VII) así como un peso de regresión aceptable (R > 0,4) y muy aceptable (R>0,7) como se presenta en la Tabla VIII con un valor RMSEA de 0,073 dentro de la normalidad (RMSEA aceptable entre 0,04 y 0,08) (Bollen,1989).

La estructura dimensional resultante de estos análisis modificó el agrupamiento teórico de algunos ítems; estructura que ha marcado el análisis de los datos y que se presenta en la Tabla IX.

Sobre la información recogida se han realizado análisis estadísticos descriptivos e inferenciales apoyados en el programa SPSS 27.0. Se ha calculado la media y la desviación típica para todas las dimensiones que integran este bloque. Se ha profundizado sobre estos resultados primarios realizando los contrastes estadísticos oportunos en función del cumplimiento de los requisitos de los mismos (T de Student, ANOVA y DMS-LSD).

De entre todas las variables consideradas en este estudio se exploran el género, el perfil (gestor, docente, prospector) y el desarrollo profesional (nivel bajo, medio y alto).

Recibido: 22/03/2023 Aceptado:22/12/2023

TABLA V. Matriz de componentes rotados

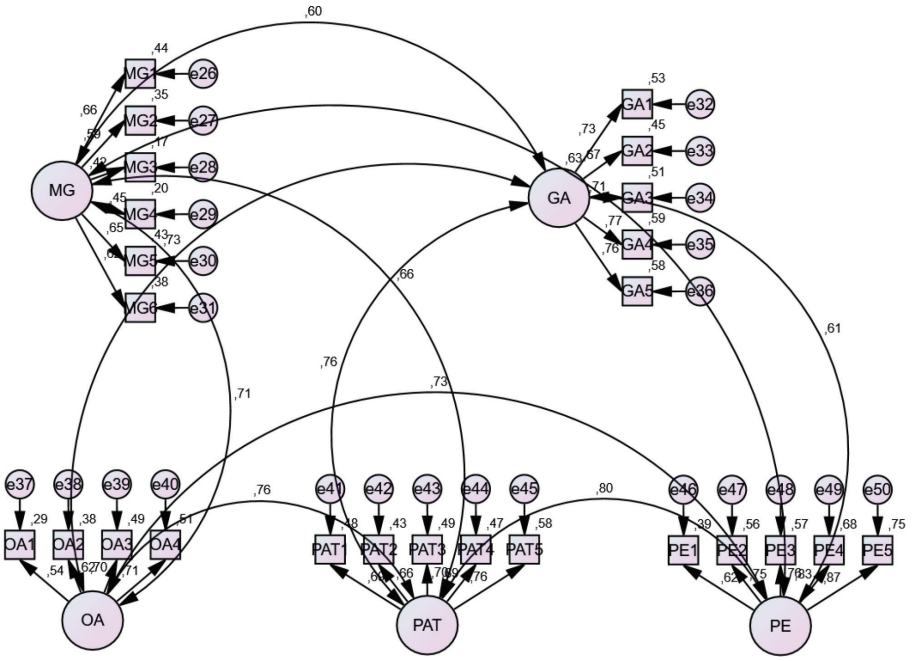
Matriz de componentes rotados (a)					
Ítem	**Componente**				
	D1(PE)	**D2 (GA)**	**D3 (PAT)**	**D4 (MG)**	**D5 (OA)**
Item1	0,07	0,11	0,09	0,58	0,50
Item2	0,09	0,09	-0,03	0,72	0,31
Item3	0,05	0,13	0,23	0,50	-0,01
Item4	0,20	0,67	0,23	0,15	0,20
Item5	0,08	0,30	0,08	0,33	0,31
Item6	0,11	0,64	0,14	0,20	0,29
Item7	0,24	0,74	0,23	0,13	-0,06
Item8	0,19	0,78	0,18	0,17	0,02
Item9	0,18	0,73	0,30	0,02	0,21
Item10	0,28	0,31	-0,11	0,37	0,38
Item11	0,17	0,41	0,42	-0,03	0,48
Item12	0,22	0,09	0,40	0,00	0,66
Item13	0,30	0,15	0,12	0,19	0,68
Item14	0,18	0,28	0,66	0,15	0,13
Item15	0,20	0,19	0,72	0,12	0,10
Item16	0,34	0,23	0,54	0,30	0,10
Item17	0,62	0,22	0,28	0,09	0,14
Item18	0,73	0,17	0,19	0,24	0,18
Item19	0,80	0,15	0,16	0,15	0,13
Item20	0,36	0,09	0,16	0,70	-0,07
Item21	0,20	0,10	0,26	0,63	0,01
Item22	0,32	0,20	0,58	0,27	0,18
Item23	0,33	0,29	0,64	0,15	0,16
Item24	0,67	0,28	0,33	0,13	0,16
Item25	0,76	0,19	0,24	0,19	0,20

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.
a. La rotación ha convergido en 8 iteraciones.

Fuente: elaboración propia.

Recibido: 22/03/2023 Aceptado:22/12/2023

FIGURA I. Análisis Factorial Confirmatorio (AFC)



Fuente: elaboración propia

TABLA VI. Valores AFC

			Estimate	S.E.	C.R.	P	Label
FMG1	<---	MG	1,00				
FMG2	<---	MG	,98	,058	16,90	***	par_11
FMG3	<---	MG	,66	,055	12,16	***	par_12
FMG4	<---	MG	,84	,064	13,15	***	par_13
FMG5	<---	MG	1,18	,069	17,16	***	par_14
FMG6	<---	MG	1,09	,067	16,41	***	par_15
FGA1	<---	GA	1,00				
FGA2	<---	GA	1,05	,046	23,06	***	par_16

(*continuada*)

TABLA VI. Valores AFC (*continuada*)

			Estimate	S.E.	C.R.	P	Label
FGA3	<---	GA	,86	,035	24,79	***	par_17
FGA4	<---	GA	1,11	,042	26,47	***	par_18
FGA5	<---	GA	1,14	,043	26,16	***	par_19
FOA1	<---	OA	1,00				
FOA2	<---	OA	,80	,050	16,08	***	par_20
FOA3	<---	OA	,91	,052	17,48	***	par_21
FOA4	<---	OA	1,09	,062	17,59	***	par_22
FPAT1	<---	PAT	1,00				
FPAT2	<---	PAT	1,00	,046	21,93	***	par_23
FPAT3	<---	PAT	1,20	,051	23,44	***	par_24
FPAT4	<---	PAT	1,09	,048	22,69	***	par_25
FPAT5	<---	PAT	1,05	,042	25,00	***	par_26
FPE1	<---	PE	1,00				
FPE2	<---	PE	1,54	,068	22,65	***	par_27
FPE3	<---	PE	1,32	,058	22,90	***	par_28
FPE4	<---	PE	1,38	,057	24,31	***	par_29
FPE5	<---	PE	1,60	,064	25,01	***	par_30

Fuente: elaboración propia.

TABLA VII. Correlaciones entre dimensiones AFC

CORRELACIONES			ESTIMACIÓN
FMG	<-->	FGA	,60
FMG	<-->	FOA	,70
FMG	<-->	FPAT	,66
FPE	<-->	FMG	,63
FGA	<-->	FOA	,73
FOA	<-->	FPAT	,76
FPE	<-->	FOA	,73
FGA	<-->	FPAT	,76
FPE	<-->	FPAT	,80
FPE	<-->	FGA	,61

Fuente: elaboración propia.

Recibido: 22/03/2023 Aceptado:22/12/2023

TABLA VIII. Regresión. AFC

Regresión			Estimación
MG1	<---	MG	,66
MG2	<---	MG	,59
MG3	<---	MG	,42
MG4	<---	MG	,44
MG5	<---	MG	,65
MG6	<---	MG	,62
GA1	<---	GA	,73
GA2	<---	GA	,67
GA3	<---	GA	,71
GA4	<---	GA	,77
GA5	<---	GA	,76
OA1	<---	OA	,54
OA2	<---	OA	,62
OA3	<---	OA	,70
OA4	<---	OA	,71
PAT1	<---	PAT	,69
PAT2	<---	PAT	,66
PAT3	<---	PAT	,70
PAT4	<---	PAT	,68
PAT5	<---	PAT	,76
PE1	<---	PE	,62
PE2	<---	PE	,74
PE3	<---	PE	,76
PE4	<---	PE	,83
PE5	<---	PE	,86

Fuente: elaboración propia.

Los análisis se han realizado desde dos perspectivas; una perspectiva general que nos ha permitido identificar en base a los estadísticos descriptivos básicos las buenas prácticas desarrolladas en los centros andaluces en lo que a la gestión de la FPD se refiere (aspectos valorados como facilitadores) y aquellas otras que presentan cierto margen de mejora (aspectos valorados como obstaculizadores). Una segunda perspectiva de

TABLA IX. Distribución de ítems por dimensiones

Dimensiones	Ítems
[D1.PE] Papel de la empresa en la FPD	PE1. La motivación del alumnado dual PE2. El conocimiento que las empresas tienen de los contenidos de los Ciclos Formativos que se imparten en los centros PE3. El compromiso de las empresas participantes con la formación del alumnado dual PE4. La coordinación de los contenidos formativos entre centro y empresa PE5. El cumplimiento del programa formativo por parte de la empresa
[D2.GA] Gestión Administrativa en el centro	GA1. El proceso seguido en el centro para la planificación de los proyectos duales GA2. La gestión administrativa de la FPD en el centro GA3. La red de contactos que el centro mantiene con las empresas del entorno GA4. El sistema implementado por el centro para la captación de empresas GA5. El procedimiento seguido en el centro para la asignación del profesor-tutor en la FPD
[D3.PAT] Profesorado y acción tutorial	PAT1. La implicación del profesorado en la FPD PAT2. La continuidad del profesorado en los proyectos duales PAT3. La formación que el profesorado del centro tiene sobre FPD PAT4. La orientación previa que recibe el alumnado para la elección de la dual PAT5. El sistema de orientación y tutorización del alumnado dual
[D4.MG] Modelo de gobernanza	MG1. Los términos y condiciones de las convocatorias anuales de la Administración educativa andaluza MG2. Los recursos asignados por la Administración al centro para la puesta en marcha de la FPD MG3. La utilidad de la tarea desempeñada por el/la prospector/a durante el primer trimestre del curso 2019/20 MG4. La existencia dentro de un mismo Ciclo de alumnado dual y no dual MG5. El papel desempeñado por las Cámaras de Comercio, Confederaciones, Asociaciones de empresarios… MG6. El papel desempeñado por organizaciones tales como fundaciones (Bertelsman, Dualiza,…) y otras instituciones que favorece la implantación de la FPD
[D5.OA] Organización académica en el centro	OA1. El tiempo asignado al/a tutor/a del centro para el seguimiento de su alumnado dual OA2. El procedimiento seguido en el centro para la selección del alumnado dual OA3. La duración de la formación en la empresa OA4. La compatibilidad de horarios entre el centro y la empresa

Fuente: elaboración propia.

análisis, centrada en las dimensiones en las que se agrupan estas prácticas, nos ha permitido valorar en qué medida las variables anteriormente referidas, condicionan la opinión del profesorado encuestado.

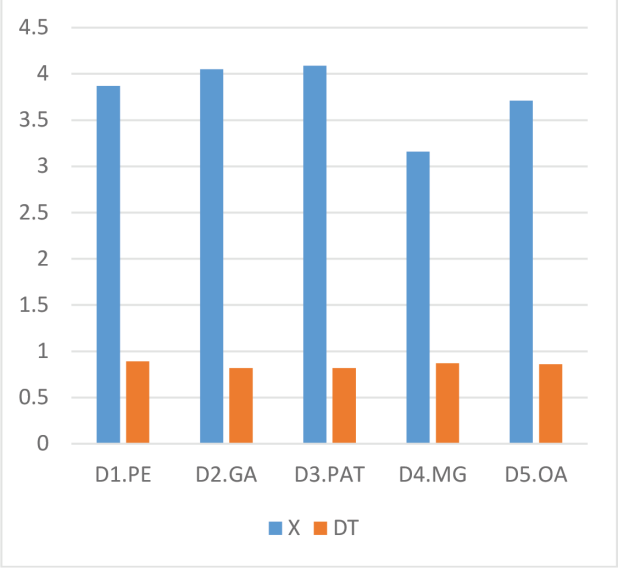
Recibido: 22/03/2023 Aceptado:22/12/2023

Resultados

Desde una perspectiva general, las prácticas que se desarrollan en los centros andaluces relacionadas con el profesorado y la acción tutorial y con la gestión administrativa que de la FPD se realiza en los mismos son las mejor valoradas y las que más favorecen la implantación de esta modalidad formativa en la Comunidad Autónoma Andaluza en opinión del colectivo docente. Las prácticas relacionadas con el modelo de gobernanza, con la organización académica y con el papel de la empresa son las que presentan cierto margen de mejora (Gráfico I).

Como se muestra en la Tabla X, de entre los aspectos y prácticas que más favorecen la implantación de esta modalidad formativa en los centros andaluces destacan la implicación del profesorado, su continuidad en los proyectos duales y la red de contactos que los centros mantienen con las empresas del entorno (valorados con un 5/5 por más

GRÁFICO I. Prácticas que favorecen u obstaculizan la implantación de la FPD en los centros andaluces por Dimensiones



*Siendo 1 "No favorece el desarrollo de la FPD en mi centro" y 5 "Favorece mucho".

Recibido: 22/03/2023 Aceptado:22/12/2023

TABLA X. Aspectos y prácticas favorecedoras / obstaculizadoras de la implantación de la FPD en los centros andaluces

Prácticas bien valoradas		**1**	**2**	**3**	**4**	**5**	**X**	**σ**
1º	GA3	1,36%	3,29%	8,30%	28,92%	58,12%	4,39	0,88
2º	PAT2	2,96%	4,81%	10,30%	22,96%	58,96%	4,30	1,03
3º	PAT1	2,32%	4,06%	11,52%	25,94%	56,16%	4,30	0,98
4º	PAT5	1,76%	4,33%	13,35%	36,24%	44,31%	4,17	0,94
5º	PE1	1,52%	5,42%	15,31%	33,07%	44,69%	4,14	0,97
6º	GA5	3,13%	2,97%	15,17%	36,67%	42,06%	4,12	0,98
7º	GA1	2,41%	5,18%	16,27%	35,81%	40,34%	4,06	0,99
8º	OA3	2,91%	4,52%	15,80%	39,84%	36,93%	4,03	0,98
9º	GA4	3,27%	5,65%	17,41%	34,00%	39,66%	4,01	1,05
10º	GA5	3,82%	5,65%	16,96%	32,85%	40,72%	4,01	1,07
11º	PE3	2,86%	6,29%	18,53%	31,62%	40,70%	4,01	1,05
12º	PE4	2,12%	6,42%	18,18%	35,55%	37,74%	4,00	1,00
13º	PAT4	4,10%	5,97%	17,54%	34,33%	38,06%	3,96	1,08
14º	MG3	5,97%	5,97%	16,16%	30,65%	41,26%	3,95	1,16
Prácticas con margen de mejora		**1**	**2**	**3**	**4**	**5**	**X**	**σ**
1º	MG2	24,75%	27,65%	26,28%	14,06%	7,26%	2,51	1,21
2º	MG4	21,19%	17,35%	24,41%	20,09%	16,95%	2,94	1,38
3º	MG5	19,04%	20,41%	26,17%	19,43%	14,94%	2,91	1,32
4º	OA1	19,31%	20,36%	19,02%	21,33%	19,99%	3,02	1,41
5º	MG6	12,18%	16,86%	25,73%	24,66%	20,57%	3,25	1,29
6º	MG1	6,55%	14,23%	32,90%	30,96%	15,36%	3,34	1,10
7º	PE2	8,31%	14,97%	23,71%	28,30%	24,71%	3,46	1,24
8º	OA4	5,77%	10,66%	23,21%	31,31%	29,05%	3,67	1,17
9º	PE5	4,36%	9,62%	20,34%	35,13%	30,55%	3,78	1,11
10º	PAT3	4,70%	10,53%	19,29%	31,17%	34,31%	3,80	1,16
11º	GA2	4,98%	8,30%	21,04%	31,83%	33,86%	3,81	1,14
*Siendo 1 "No favorece el desarrollo de la FP Dual en mi centro" y 5 "Favorece mucho"								

Fuente: elaboración propia

del 50% de los encuestados); a estos les siguen el sistema de orientación y tutorización del alumnado dual, el procedimiento seguido en el centro para la selección de este alumnado y para la planificación de los

Recibido: 22/03/2023 Aceptado:22/12/2023

proyectos duales y la duración de la formación en la empresa (valorados con una media de 4/5 y 5/5 por más del 75% de los encuestados). Para este profesorado la motivación del alumnado constituye un aspecto que facilita la implantación de esta modalidad formativa (4,14).

De entre las prácticas con mayor margen de mejora destacan (con valores medios situados en el intervalo 2-3/5 y desviaciones mayores a 1), los recursos asignados por la Administración al centro, la existencia dentro de un mismo Ciclo de alumnado dual y no dual, el papel desempeñado por las Cámaras de Comercio, Confederaciones y Asociaciones de empresarios y el tiempo asignado a la persona tutora del centro para el seguimiento de su alumnado dual. A este grupo también pertenecen otras prácticas cuyo margen de mejora es menor (valores medios situados en el intervalo 3-4/5 y desviaciones mayores a 1).

Conocer en qué medida el género, la experiencia y conocimiento que de esta modalidad formativa tiene el colectivo docente (Desarrollo profesional) y el perfil profesional (Gestor – Docente – Prospector/a) condicionan la opinión que de estas prácticas tiene el profesorado encuestado constituye el objeto de otros análisis realizados y cuyos resultados y discusión mostramos a continuación.

Estas variables no parecen condicionar la valoración general que las personas encuestadas hacen de las prácticas desarrolladas en los centros; así pues, las prácticas mejor valoradas tienen que ver con la gestión administrativa de la dual en el centro y con el profesorado y la acción tutorial y las que más margen de mejora presentan están relacionadas con el modelo de gobernanza; no obstante, los análisis realizados han permitido identificar algunas diferencias significativas en los niveles de concreción de dichas variables (Tabla XI).

La T de Student realizada nos permite identificar que existen diferencias significativas ($p < 0,05$) entre las opiniones de los profesores y profesoras en cuatro de las cinco dimensiones objeto de estudio; siendo la dimensión relativa a la Gestión administrativa de la FPD en el centro la única que no muestra diferencias en función del género.

El análisis ANOVA realizado, muestra que existen diferencias significativas en las valoraciones de las cinco dimensiones analizadas atendiendo al perfil del profesorado encuestado: gestión, docencia y prospección ($p <0,05$).

Como se muestra en la Tabla XII, el profesorado con perfil de gestor, muestra las valoraciones más altas frente a aquel profesorado prospector, cuya valoración es la más baja.

TABLA XI. Puntuaciones medias de cada dimensión según variables de análisis

Dimensión		**PE**		**GA**		**PAT**		**MG**		**OA**	
Variables		***X***	***σ***	***X***	***σ***	***X***	***σ***	***X***	***σ***	***X***	***σ***
Género	H	3,80	0,91	4,04	0,82	4,01	0,85	3,10	0,86	3,65	0,88
	M	3,97	0,85	4,06	0,82	4,18	0,78	3,23	0,88	3,78	0,85
Perfil Profesional	D	3,75	0,95	4,03	0,86	4,09	0,86	3,07	0,94	3,68	0,94
	P	3,83	0,92	3,85	0,91	3,85	0,92	3,33	0,81	3,56	0,94
	G	4,02	0,78	4,17	0,71	4,20	0,71	3,15	0,82	3,80	0,73
Desarrollo Profesional	B	3,59	1,03	3,89	0,90	3,86	0,97	3,00	0,94	3,58	1,01
	M	3,88	0,85	4,11	0,75	4,18	0,72	3,18	0,86	3,75	0,83
	A	4,06	0,76	4,33	0,78	4,37	0,68	3,18	0,92	3,89	0,68

Fuente: elaboración propia.

TABLA XII. Puntuación media según perfil profesional

DIMENSIONES	**Perfil profesional**					
	Docente		**Prospector**		**Gestor**	
	Media	**Desviación típica**	**Media**	**Desviación típica**	**Media**	**Desviación típica**
PE	3,75	0,95	3,83	0,92	4,02	0,78
GA	4,03	0,86	3,85	0,91	4,17	0,71
PAT	4,09	0,86	3,85	0,92	4,20	0,71
MG	3,07	0,94	3,33	0,81	3,15	0,82
OA	3,68	0,94	3,56	0,94	3,80	0,73

Fuente: elaboración propia.

Tras realizar el cálculo de la diferencia mínima significativa (análisis DMS) entre las distintas variables, destacamos que estas diferencias están presentes en los aspectos que integran las dimensiones "Gestión administrativa de la FPD en el centro" (GA) y "Profesorado y acción tutorial" (PAT) para los tres perfiles (p=entre 0,00 y 0,02). Las diferencias en los aspectos de la dimensión "Modelo de gobernanza" (MG) se dan entre las opiniones del profesorado encuestado con perfil docente y las del perfil de prospección (p=0,00), y entre este y el de gestión (p=0,01);

Recibido: 22/03/2023 Aceptado:22/12/2023

las diferencias en las valoraciones de los distintos aspectos que integran las dimensiones relativas a la "Organización académica" (OA) y al "Papel de la empresa" (PE) las marcan los gestores, cuyas opiniones difieren de las de los docentes y prospectores (p entre 0,00 y 0,02).

Aunque para todas las personas encuestadas no resulta suficiente el tiempo que el centro asigna al/la tutor/a para el seguimiento del alumnado dual, esta valoración es significativamente más baja para los que conocen y desempeñan esta tarea (2,87/5; p=,008).

En el caso de las personas encuestadas con responsabilidad en la coordinación de proyectos duales, los aspectos explorados (por las características de su tarea) están relacionados con la gestión administrativa y académica de la FPD en el centro y con el apoyo de la Administración educativa andaluza. En ocho de los diez aspectos explorados (Modelo de Gobernanza, ítems 1 y 4; Gestión Administrativa, ítems 1, 2 y 3; Organización Académica, ítems 2, 3 y 4), las valoraciones medias de los coordinadores son mayores que la de los encuestados sin esta experiencia; según los datos del ANOVA realizado, estas diferencias son significativas (con un valor de p=,000) en las opiniones sobre el proceso seguido en el centro para la planificación de los proyectos duales (GA1) y para la selección del alumnado dual (OA2) y la duración de la formación en la empresa (OA3). También destacan significativamente (con un valor de $p \leq 0,05$) la valoración positiva que este colectivo hace de la red de contactos que el centro mantiene con las empresas (GA3), de la existencia dentro del mismo ciclo de alumnado dual y no dual (MG4) y de la compatibilidad de horarios entre el centro y la empresa (OA4). De las prácticas valoradas por este grupo mejorables (con valores medios por debajo, de los encuestados que nunca han participado en la coordinación de proyectos) destaca el tiempo asignado al/la tutor/a del centro para el seguimiento de su alumnado dual (OA1) con diferencias estadísticamente significativas (p=,004).

En relación al índice de desarrollo profesional, hay que destacar que existen diferencias significativas en todas las dimensiones tomando como referente el valor de $p \leq 0,05$: Papel de la empresa: p=0,000; Gestión administrativa: p=0,000; Profesorado y acción tutorial: p=0,000; Modelo de gobernanza: p=0,035; Organización académica: p=0,006.

En términos generales, las opiniones medias del profesorado con gran experiencia en FPD (índice alto) son más altas que las del resto en todas las dimensiones menos en Modelo de Gobernanza (con una media entre

3,19 y 4,37), igualándose estas valoraciones con las del profesorado con un índice de desarrollo profesional medio en la dimensión relativa al Modelo de Gobernanza (media 3,18).

Destaca que las opiniones medias del profesorado con menos experiencia (índice bajo) son más bajas que las del resto (media que oscila entre 3 y 3,89), en todas las dimensiones. El análisis DMS realizado identifica que las diferencias mínimas significativas en los distintos aspectos de las dimensiones "Gestión administrativa de la FPD en el centro (GA)" y "Profesorado y acción tutorial (PAT)" se dan entre todos los colectivos ($p < 0,05$). Las diferencias en los aspectos de la dimensión "Modelo de gobernanza (MG)" se dan entre las opiniones del colectivo docente con un nivel bajo y medio de experiencia en FPD. Las diferencias en las valoraciones de los aspectos de la dimensión relacionada con la "Organización académica (OA)" se dan entre el profesorado con nivel bajo y medio, bajo y alto y alto y bajo (p entre 0,00 y 0,05). En la dimensión "Papel de la empresa en la formación dual", son las opiniones del profesorado con un desarrollo profesional bajo las que se diferencian de las de los demás (nivel medio y alto).

Discusión

Como muestran los análisis presentados anteriormente, la implicación y continuidad del profesorado en los proyectos duales constituye uno de los aspectos más favorecedores de la implantación de esta modalidad formativa en los centros andaluces. Sin embargo, el profesorado parece no sentirse del todo preparado (3,5/5 y DT 1,2), para asumir su tarea en esta modalidad de FP, lo que nos informa de un cambio en la respuesta disciplinada y obediente que este colectivo mostró al inicio de la implantación de esta modalidad de FP (Marhuenda, et al., 2017); la demanda del profesorado en este sentido lo sitúa en una posición mucho más proactiva.

El esfuerzo organizacional académico y administrativo que realizan los centros para la puesta en marcha de la formación en alternancia es grande y su eficacia muy bien valorada por el profesorado; en este sentido destacan las redes de contactos que mantienen los centros con las empresas del entorno y los procedimientos seguidos por estos para la selección del alumnado, la planificación de los proyectos duales y la asignación del profesorado-tutor.

Recibido: 22/03/2023 Aceptado:22/12/2023

En relación al papel de la empresa, y sin perder de vista que la gran motivación con la que el alumnado acude a la misma constituye una garantía del éxito del proceso formativo desarrollado en ella, destaca el compromiso de esta con la formación y su coordinación con el centro. Sin embargo, tal y como se destaca en estudios previos (Barrientos, et al., 2019; Pozo-Llorente & Poza-Vilches, 2020; Pozo-Llorente et al., 2021, entre otros), en opinión del profesorado, debe mejorar el conocimiento que las empresas tienen de los contenidos de estos Ciclos de FP y por lo tanto su grado de cumplimiento del programa formativo.

La complejidad de los sistemas duales de FP, debido a la variedad de agentes implicados (administraciones públicas, centros, empresas, agentes sociales, etc.), ha convertido su gobernanza en una pieza clave. Como ya señaló Oliver (2010) en relación a la FP, su éxito y eficacia depende, en buena medida, del desarrollo de unas estructuras de gobernanza que favorezcan la coordinación e integración de los distintos actores que participan. En línea con lo evidenciado por el estudio realizado por Sanz de Miguel (2017), en el sistema dual andaluz las prácticas relacionadas con el modelo de gobernanza impuesto por las directrices de las Administraciones y que condiciona la organización en los centros son las que tienen mayor margen de mejora. En este sentido se valora como insuficiente el tiempo que tiene el profesorado tutor para el seguimiento de su alumnado dual y la compatibilidad de horarios entre el centro y la empresa. Este aspecto constituye un hándicap para empresas colaboradoras cuyo horario no puede ajustarse al lectivo ya que deben responder a situaciones excepcionales que demandan un horario especial; situaciones que también constituyen escenarios formativos para el estudiantado. Es necesario destacar que este aspecto se ha flexibilizado en la última normativa regulatoria de las convocatorias anuales de proyectos de FPD en Andalucía[6].

La única práctica relacionada con el modelo de gobernanza bien valorada por el profesorado es la relacionada con la tarea desempeñada por el/la prospector/a. Esta figura fue introducida por la Administración andaluza en el sistema dual en 2019 como respuesta a la demanda de mayor liderazgo y dotación de recursos realizada desde distintos ámbitos

[6]. https://www.juntadeandalucia.es/educacion/portals/web/inspeccion-educativa/normativa/-/normativas/detalle/orden-de-20-de-enero-de-2023-por-la-que-se-convocan-proyectos-de-formacion-profesional-dual-para-el-curso-academico

y con el propósito de organizar y fomentar en los centros públicos la prospección de empresas para la captación de plazas formativas para el alumnado de dual.

Es necesario destacar que en opinión del profesorado, en Andalucía la existencia dentro de un mismo ciclo y por lo tanto en un mismo grupo, de alumnado dual y no dual constituye un aspecto condicionador de la implantación exitosa de estas enseñanzas, valoración esta que no coincide con lo señalado por Pineda et al. (2019) en su estudio multirregional en el que destacan que "combinar formación dual y no dual de manera simultánea no afecta en general a los factores para la implementación" (p. 35). Posiblemente la elevada ratio profesorado/alumnado en los Ciclos de FP en los centros andaluces dificulte la labor docente del profesorado con alumnado dual y no dual en una misma aula.

El género, la experiencia en gestión y el conocimiento del sistema dual son variables que condicionan la opinión del profesorado sobre el proceso de implantación de estas enseñanzas en los centros andaluces. En términos generales, las profesoras y el profesorado con experiencia en gestión (sobre todo en coordinación de proyectos duales) muestran las valoraciones más altas. El profesorado encuestado que ha desarrollado tareas de prospección y el que tiene un nivel bajo de desarrollo profesional es el que realiza las valoraciones más bajas. En el caso de los/as prospectores/as posiblemente tenga que ver el gran esfuerzo que realizan hasta conseguir cerrar convenios de colaboración con cierta permanencia con las empresas y entidades colaboradoras; en el caso del profesorado con nivel bajo de desarrollo profesional, el desconocimiento del sistema dual puede ser el condicionante de su opinión (Pozo-Llorente et al., 2021).

Conclusiones

Este estudio se ha abordado desde la premisa de que el análisis específico de la singularidad local (regional) nos permitirá avanzar hacia un sistema dual más eficaz, eficiente y sostenible a nivel global. Son muchas las voces (políticas, técnicas y científicas) que insisten en que, junto con la voluntad política y el tejido productivo, el éxito de este modelo reside en el profesorado y en las empresas implicadas. En este sentido insisten Ramasamy, et al. (2021) al destacar al profesorado como elemento clave

Recibido: 22/03/2023 Aceptado:22/12/2023

en la calidad de esta formación. Como señalan Pineda et al. (2018), para conseguir una mayor implicación de estos colectivos es necesario que el profesorado entienda esta modalidad de FP y que las propuestas tengan en cuenta las motivaciones y necesidades de las empresas, sin perder de vista que el objetivo último es la formación del alumnado, destacando así la importancia del trabajo coordinado de los centros educativos y las empresas y el papel de liderazgo que debe desempeñar la Administración.

Este trabajo nos ha permitido entender algo más sobre la FPD a partir de la opinión que de su implementación en los centros andaluces tiene el profesorado implicado, así como ahondar en el estado del debate que este modelo ha generado en la comunidad educativa andaluza y del que las directrices administrativas anuales se han hecho eco.

Este estudio tiene limitaciones en términos de posibles sesgos provocados por la situación de estrés vivida en la etapa dura de pandemia en la que se desarrolló parte del mismo. El reajuste del trabajo de campo a esta situación, su adaptación a los ritmos de los centros, el carácter de la información buscada y los procesos de triangulación desarrollados, fueron las estrategias puestas en marcha para minimizar el impacto de la situación de excepcionalidad en las opiniones recogidas. Otra limitación de este estudio ha sido la falta de exploración de la opinión del profesorado en base a la familia profesional en la que imparte docencia.

Los cambios que en la práctica de las aulas duales andaluzas está provocando la nueva Ley de FP (B.O.E., 2022) y el análisis del potencial del sistema dual defendido en la misma para responder a la situación profesional en el contexto andaluz, constituyen aspectos de especial interés y líneas de estudio futuras.

Finalizamos este trabajo insistiendo en que hay que seguir apostando por el modelo dual y para ello hay que analizarlo desde el prisma de sus protagonistas diarios.

Financiación

Proyecto Marco de Financiación: La Formación Profesional Dual en Andalucía: Situación actual y retos de futuro (2019-2021), financiado por la Fundación Bankia para la Formación Profesional Dual. Investigadores coordinadores: Mª Teresa Pozo Llorente y José Manuel Aguayo Moral.

Referencias bibliográficas

Aguilar, M.C. (2015). El sistema "alternativo" de formación profesional dual para el empleo en España: perspectivas de evolución. *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo*, *3*(4),1-26.

Barrientos, D., Martín Artiles, A., Lope, A., & Carrasquer, P. (2019). La FP Dual y la transición de los jóvenes al mercado de trabajo: la visión de los agentes sociales. Anuario de trabajo y relaciones laborales. *Transiciones al empleo I*(6), 75–94. http://dx.doi.org/10.5565/rev/aiet.85

Bentolila, S., & Jansen, M. (2019). La implantación de la FP Dual en España: La experiencia de Madrid. *Información Comercial Española, ICE: Revista de Economía, 910,* 65-7922. https://doi.org/10.32796/ice.2019.910.6923

B.O.E. (2012). *Real Decreto 1529/2012 de 8 de noviembre, por el que se desarrolla el contrato para la formación y el aprendizaje y se establecen las bases de la formación profesional dual*. https://www.boe.es/buscar/doc.php?id=BOE-A-2012-13846

B.O.E. (2022). *Ley Orgánica 3/2022, de 31 de marzo, de ordenación e integración de la Formación Profesional*. https://www.boe. es/eli/es/lo/2022/03/31/3/dof/spa/

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons. Inc. https://doi.org/10.1002/9781118619179

CEDEFOP- European Centre for the Development of Vocational Training (2020). *Vocational education and training in Europe, 1995-2035. Scenarios for European vocational education and training in the 21st century*. Publications Office of the European Union.

Comisión Europea (2020). Eurostat. https://bit.ly/2TLfNWy

Echeverría, B. (2016a). Primer Trienio de Formación Profesional Dual. *Educaweb*. https://www.educaweb.com/noticia/2016/01/12/primer-trienio-formacion-profesional-dual-9201/

Echeverría, B. (2016b). Transferencia del Sistema de FP Dual a España. *Revista de Investigación Educativa*, *34*(2), 295-314. http://dx.doi.org/10.6018/rie.34.2.249341

Echeverría, B., & Martínez Clares, P. (2018). Estrategias de Mejora en la Implantación de la Formación Profesional Dual en España. *EKONOMIAZ. Revista. Vasca de Economía*, *94*, 178–203.

Recibido: 22/03/2023 Aceptado:22/12/2023

Echeverría, B., & Martínez Clares, P. (2021). Perspectiva Internacional de la investigación sobre Formación Profesional en España. *Educación XX1, 24*(2), 231-256. https://doi.org/10.5944/educXX1.28178

Emmenegger, P., & Seitzl, L. (2020). Social partner involvement in collective skill formation governance. A comparison of Austria, Denmark, Germany, the Netherlands and Switzerland. *European Review of Labour and Research, 26*(1), 27-42.

Igartua, M.T. (2017). La formación dual en el sistema educativo: balance y propuestas de mejora. *Temas laborales: Revista andaluza de trabajo y bienestar social, 137*, 91-125.

Ishar, M.I.M., Derahman, W.M.F.W., & Kamin, Y. (2020) Practices and planning of ministries and institutions of technical and vocational educational training (TVET) in facing the industrial revolution 4.0 (IR4.0). *Malaysian Journal of Social Sciences and Humanities (MJSSH), 5*(3), 47-50. http://dx.doi.org/10.47405/mjssh.v5i3.374

Marhuenda, F., Gisvert, M.J., Palomares, D., & Vila, J. (2017). Con *d de dual*: investigación sobre la implantación del sistema dual en la Formación Profesional en España. *Educar*, *53*(2), 285-307 https://doi.org/10.5565/rev/educar.821

Martín Artiles, A., Lope, A., Barrientos, D., Moles, B., & Carrasquer, P. (2020) The rhetoric of europeanisation of Dual Vocational Education and Training in Spain. *Transfer, 26*, 1, 73–90. https://doi.org/10.1177/1024258919896901

Martínez Morales, I., y Marhuenda, F. (2020). Vocational education and training in Spain: steady improvement and increasing value. *Journal of Vocational Education & Training*, 1-19. https://doi.org/10.1080/13636820.2020.1729840

Ministerio de Educación y Formación Profesional de España (MEFP) (2020a). *I Plan Estratégico de Formación profesional del Sistema Educativo 2019-2022*. Recuperado en: https://www.todofp.es/dam/jcr:163978c0-a214-471e-868d-82862b5a3aa3/plan-estrategico--enero-2020.pdf

Ministerio de Educación y Formación Profesional de España (MEFP) (2020b). *Plan de Modernización de la Formación Profesional 2020*. Recuperado en: https://www.todofp.es/dam/jcr:5d43ab06-7cdf-4db6-a95c-b97b4a0e1b74/220720-plan-modernizacion-fp.pdf

Ministerio de Educación y Formación Profesional de España (MEFP) (2020c). *Panorama de la Educación: Indicadores de la OCDE 2020. Informe español*. Madrid: Instituto Nacional de Evaluación Educativa. Recuperado en: https://www.educacionyfp.gob.es/inee/indicadores/indicadores-internacionales/ocde.html

Navarro, M. (Dir.) (2019). *El sistema de Formación Profesional de La Rioja*. Madrid: Fundación Bankia por la Formación Dual.

Oliver, D. (2010): Complexity in Vocational Education and Training Governance, *Research in Comparative and International Education, 5*(3), 261-273. http://dx.doi.org/10.2304/rcie.2010.5.3.261

Pineda, P., Ciraso, A., & Arnau, L. (2019). La FP dual desde la perspectiva del profesorado: elementos que condicionan su implementación en los centros. *Educación XX1*, *22*(1), 15-43, https://doi.org/10.5944/educXX1.21242

Pineda, P., Fernández, M., Espona, B., & Grollman, P.C. (2018). Formación profesional dual. Las motivaciones de las empresas del sector de la automoción en España para involucrarse en formación profesional dual. *Revista de educación*, *382*, 33-56. https://doi.org/10.4438/1988-592X-RE-2018-382-391

Pozo-Llorente, M.T., Aguayo-Moral, J.M., Poza-Vilches, M.F., Moreno-Verdejo, A., Navarro-Ruiz, M.A. & del Río-Cabeza, A.I. (2021). *Diagnóstico de la Formación Profesional Dual en Andalucía y propuestas de mejora (2013-2020)*. CaixaBank Dualiza. https://www.caixabankdualiza.es/

Pozo-Llorente, M.T. & Poza-Vilches, M.F. (2020). Conditioning factors of sustainability of dual vocational educational training in Andalusia (Spain): Case study of three educational centres. *Sustainability*, *12*(22), 9356.

Ramasamy, M., Regel, J., Sharma, H., Rajagopalan, A., & Pilz, M. (2021). Measuring Quality in Indian VET Institutions: Development Steps Towards a Framework Adapted to the National Context. *International Journal for Research in Vocational Education and Training (IJRVET), 8*(4), 44–66. https://doi.org/10.13152/IJRVET.8.4.3

Sanz de Miguel, P. (2017). Gobernanza de la formación profesional dual española: entre la descoordinación y la falta de objetivos. *Revista Española De Educación Comparada*, *30*, 77–98. https://doi.org/10.5944/reec.30.2017.18705

Šćepanović, V.M, & Martín Artiles, A. (2020). Dual training in Europe: a policy fad or a policy turn? *Transfer: European Review of Labour and Research*, *26*(1), 15-26. https://doi.org/10.1177/1024258920902584

Stephen, O. O. (2020). *Information Technology, Technical Vocational Education in Developing Workforce Towards Globalization. The Roles of Technology and Globalization in Educational Transformation*. IGI Global; pp. 80-97.

Recibido: 22/03/2023 Aceptado:22/12/2023

Vaganova, I., Odarich, N., Rudakova, A., Volkova, S., & Bulaeva, N. (2020). Level differentiation technology in vocational education. *Revista Amazonía investiga, 9*(28), 42-49 https://doi.org/10.34069/ai/2020.28.04.5

Vila, J.J., & Chisvert, M.J. (2018). *Luces y sombras de la Formación Profesional Dual en el sistema educativo español*. Tirant Humanidades.

Virgós, V., & Burguera, J. (2020). Evaluación del Proceso formativo de tutores de empresa en la Formación Profesional Dual. *Education in the Knnowledge Society, 21*, 1-11. https://doi.org/10.14201/eks.22992

Virgós, V. Burguera, J., & Pérez, M (2022). La Formación Profesional Dual en la empresa desde la perspectiva de sus protagonistas. *Revista Complutense de Educación, 33*(1), 27-39 https://doi.org/10.5209/rced.70992

Información de contacto: María Teresa Pozo Llorente. Universidad de Granada, Campus Universitario de Cartuja, Facultad de Ciencias de la Educación. 18071. Granada. E-mail: mtpozo@ugr.es

Anexo I

Dimensiones teóricas e ítems asociados derivados de la consulta a personas expertas en FPD en Andalucía

BLOQUE I. ASPECTOS QUE FAVORECEN LA IMPLANTACIÓN DE LA FPD EN ANDALUCÍA

NORMATIVA
Los términos y condiciones de las convocatorias anuales de la Administración Educativa andaluza
APOYO ADMINISTRACIÓN
Los recursos asignados por la Administración al centre para la puesta en marcha de la FPD
La utilidad de la tarea desempeñada por el/la prospector/a durante el primer trimestre del curso 2019/20
DISEÑO Y OFERTA
El proceso seguido en el centro para la planificación de los proyectos duales
La existencia dentro de un mismo Ciclo de alumnado dual y no dual
GESTIÓN ADMINISTRATIVA Y EDUCATIVA
La gestión administrativa de la FPD en el centro
La red de contactos que el centro mantiene con las empresas del entorno
El sistema implementado por el centro para la captación de empresas
El procedimiento seguido en el centro para la asignación del profesor-tutor en la FPD
El tiempo asignado al/a tutor/a del centro para el seguimiento de su alumnado dual
El procedimiento seguido en el centro para la selección del alumnado dual
La duración de la formación en la empresa
La compatibilidad de horarios entre el centro y la empresa
AGENTES IMPLICADOS
La implicación del profesorado en la FPD
La continuidad del profesorado en los proyectos duales
La formatión que el profesorado del centro tiene sob re FPD
La motivación del alumnado dual
El conocimiento que las empresas tienen de los contenidos de los Ciclos formativos que se imparten en los centros
El compromiso de las empresas participantes con la formación del alumnado dual
El papel desempeñado por las Cámaras de Comercio, Confederacies, Asociaciones de empresarios,...
El papel desempeñado por organizaciones tales como la Fundación Bertelsmann, Andalucía Emprende, la Fundación Bankia por la FPD, la Alianza por la FPD...

Recibido: 22/03/2023 Aceptado:22/12/2023

CAUDAD DE LA FORMACIÓN
La orientación previa que recibe el alumnado para la elección de la dual
El sistema de orientación y tutorización del alumnado dual
La coordinación de los contenidos formativos entre centro y empresa
El cumplimiento del programa formativo por parte de la empresa

BLOQUE II. ASPECTOS RELATIVOS A FPD OBJETO DE REVISÓN PARA MEJORAR LA CALI DAD DE LA FPD EN AN DALUCIA

N0RMATIVA
La normativa reguladora de la FPD
Las funciones y responsabilidades de todos los agentes implicados en la FPD establecidas en la normativa
APOYO ADMINISTRACIÓN
La difusión que la Administración realiza de la FPD entre los centros
La difusión que la Administración realiza de la FPD entre las empresas
El reconocimiento de la Administración al profesorado implicado
Las ayudas (beca/remuneración) al alumnado dual
Los mecanismos de seguimiento establecidos por la Administración para garantizar el buen desarrollo de los proyectos duales
DISEÑO Y OFERTA
El proceso establecido en la regulación andaluza para la determinación de la oferta de FPD
El ajuste de la oferta dual a las demandas del tejido productivo del territorio
Los criterios aplicados por la Administración para la evaluation de los proyectos
La organización de la alternancia (horas de formación en el centro y en la empresa)
Las estrategias implementadas para asegurar la presencia paritaria de alumnos y alumnas en todos los Ciclos duales
GESTIÓN ADMINISTRATIVA Y ACADÉMICA
El modelo de gestión que conlleva la puesta en marcha de la FPD en un centro
El sistema establecido para que el alumnado acceda a las FPD sin estar diferenciado de la FP presencial
La coordinación centro-empresa
Los mecanismos puestos en marcha para asegurar la admisión del alumnado más vulnerable en esta tipología de FP

Recibido: 22/03/2023 Aceptado:22/12/2023

INFRAESTRUCTURAS Y RECURSOS
Las infraestructuras y equipamientos de los que disponen los centros para el desarrollo de los proyectos duales
Las infraestructuras y equipamientos de los que disponen las empresas participantes para el desarrollo de los proyectos duales
CALIDAD DE LA FORMACIÓN
El currículum, las competencias (conocimientos, habilidades, destrezas, valores) y los resultados de aprendizaje establecidos para la FPD
La plataforma MOODLE de trabajo colaborativo entre el profesorado y los coordinadores/as del FPD
Los mecanismos de comunicación con las empresas
Las metodologías docentes ligadas a la FPD
La cualificación didáctica y pedagógica del/la tutor/a de empresa
El sistema de evaluación de los aprendizajes del alumnado dual
La rotación del alumnado por distintas empresas
El nivel real de cualificación profesional que adquiere el alumnado dual

BLOQUE III. SITUACIONES QUE PUEDEN SUPONER RIESGOS REALES PARA EL ÉXITO DE LA FPD EN ANDALUCIA

Que la complejidad que la gestión de la FPD implica para un centro suponga la pérdida de interés de los centros por esta tipología de FP
El nivel de participación en la FPD del profesorado considerando el esfuerzo que ello conlleva
El grado de interés de las empresas en el perfil competencial de los egresados de FPD
El ajuste del alumnado que finaliza la FPD a los puestos de trabajo ofertados por las empresas.
Que la reducción de horas de docencia presenciales que implica la FPD suponga, en algun caso, la disminución de la plantilla de profesorado en los centros
El grado de estabilidad en la relacion entre centro y empresa
La especialización de la formación recibida en la empresa podría implicar una disminución de las posibilidades de inserción laboral del alumnado dual
El grado de estabilidad del profesorado participante
La tasa de abandono escolar del alumnado dual
OTROS:

Recibido: 22/03/2023 Aceptado:22/12/2023

BLOQUE IV. GRADO DE SATISFACCIÓN GENERAL CON LA FPD EN ANDALUCÍA

La implantación de los proyectos de FPD en los que participa
Implantación general de la FPD en la Comunidad Andaluza
La calidad de la formación profesional dual en Andaluda

BLOQUE V. PROSPECCIÓN

La regulación de la figura del/la prospector/a es una demanda de los centros con FP
El alcance y extensión de la FPD, es decir, su capacidad para llegar a todos los territories es adecuado
Las funciónes asignadas al prospector/a por la Administración están definidas daramente
La remuneración establecióa por parte de la Administración para la figura del/la prospector/a es adecuada
La tarea desempeñada por los/as prospectores/as es básica para asegurar la partidpación de las empresas del entomo en la FPD
Para las empresas el/la prospector/a constituye el contacto con los centros educativos
El número de prospectores/as en mi entorno ha sido suficiente para alcanzar nuestro objetivo
La duración establecida por la Administración para la prospección de empresas ha sido suficiente
La función del/la prospector/a es entendida por todo el profesorado de FP de los centros
Me he sentido muy motivado por el reconocimiento de mi trabajo como prospector/a por parte de mis compañeros/as del centro
He desempeñado mis tareas como prospectores/s en permanente coordinación con mi centro
En la mayoría de las ocasiones me he sentido seguro/a realizando mi labor como prospector/a
El trabajo que he realizado como prospectores/as ha mejorado la red de contactos de los centros con las empresas del entorno.
El número de plazas comprometidas por las empresas para la FPD a partir de mi tarea como prospectos/a son suficientes para mejorar el desarrollo de la misma.
En términos generates las empresas se muestran dispuestas a becar al alumnado dual
En términos generates las empresas conocen los procedimientos para becar al alumnado dual
He alcanzado los objetivos que me marqué en mi función como prospector/a
Me gustaría repetir como prospector/a en cursos próximos
Estoy muy satisfecho con mi labor desempeñada como prospector/a

Recibido: 22/03/2023 Aceptado:22/12/2023

Formación para el desarrollo profesional docente en escuelas: revisión sistemática

Training for teacher professional development in schools: systematic review

https://doi.org/10.4438/1988-592X-RE-2024-406-643

Celia Camili Trujillo
https://orcid.org/0000-0001-7181-0068
Universidad Complutense de Madrid

Lorena Gil Pastor
https://orcid.org/0000-0001-7055-9908
Universidad Complutense de Madrid

Sara Serrano Díaz
https://orcid.org/0000-0002-7048-646x
Universidad Complutense de Madrid

Luis Aimá González
https://orcid.org/0000-0001-6557-9066
Universidad Complutense de Madrid

Resumen

Los retos educativos del siglo XXI exigen una formación para el desarrollo profesional (DP) que atienda a las demandas de los estudiantes, avances en el conocimiento de la materia y el desarrollo de competencias para una enseñanza de calidad. El objetivo del estudio es realizar una revisión sistemática de artículos científicos para analizar la formación que reciben docentes en ejercicio de educación primaria, secundaria y bachillerato para su DP, qué caracteriza a esos modelos de formación y cuál ha sido su impacto. Las bases de datos consultadas fueron SCOPUS, Web of Science, Education Resources Information

Recibido: 26/03/2023 Aceptado: 03/05/2024

Center y Dialnet. De 586 documentos encontrados finalmente 68 cumplieron con los criterios de inclusión. Las variables extrínsecas (año de publicación y procedencia de los autores) y metodológicas (muestra y método de investigación empleado) se analizaron con estadística descriptiva (frecuencias y porcentajes) y las variables sustantivas (en qué se forman los docentes, características de los modelos de formación recibida e impacto) a través de un análisis temático cualitativo. La formación se ha incrementado entre 2019-2021 debido a la pandemia COVID-19, los docentes de primaria son los que mayor formación han recibido, predomina la formación en competencias digitales y para la atención a la diversidad siendo el modelo de formación colaborativo el más trabajado. El impacto de la formación en el DP es variado y no siempre se valora por las escuelas como suficiente o adecuado. Entre los factores que disminuyen la calidad de la formación se encuentran: dificultades en la planificación del programa, incompatibilidad con el calendario laboral, carga de trabajo y falta de motivación. El estudio pretende contribuir a la toma de decisiones basada en la evidencia para el DP de docentes en ejercicio de escuelas. Se insta a una formación sistémica y multifacética en donde participen grupos interconectados de profesores, administradores y directivos.

Palabras clave: desarrollo profesional, formación, revisión sistemática, educación básica, educación obligatoria.

Abstract

The educational challenges of the 21st century require training for professional development (PD) that meets the demands of students, advances in knowledge of the subject and the development of skills for quality teaching. The objective of the study is to carry out a systematic review of scientific articles to analyze the training received by practicing teachers of primary, secondary and high school for their PD, what characterizes these training models and what their impact has been. The databases consulted were SCOPUS, Web of Science, Education Resources Information Center and Dialnet. Of 586 documents finally found, 68 met the inclusion criteria. The extrinsic variables (year of publication and origin of the authors) and methodological (sample and research method used) were analyzed with descriptive statistics (frequencies and percentages) and the substantive variables (what teachers are trained in, characteristics of the teaching models). training received and impact) through a qualitative thematic analysis. Training has increased between 2019-2021 due to the COVID-19 pandemic, primary school teachers are the ones who have received the most training, training in digital skills and attention to diversity predominates, with the collaborative training model being the most worked up. The impact of training in PD is varied and is not always valued by schools as sufficient or adequate. Among the factors that reduce the quality of training are: difficulties

Recibido: 26/03/2023 Aceptado: 03/05/2024

in planning the program, incompatibility with the work schedule, workload and lack of motivation. The study aims to contribute to evidence-based decision making for the PD of practicing school teachers. A systemic and multifaceted training is urged in which interconnected groups of teachers, administrators and managers participate.

Keywords: professional development, training, systematic review, basic education, obligatory education.

Introducción

Hoy en día se requieren docentes formados en competencias, estrategias y habilidades para dar respuesta a la diversidad del alumnado y en el uso cada vez mayor de las tecnologías de la información y la comunicación (TIC), así como en los avances en el conocimiento de la materia que imparten. Un aprendizaje permanente para el desarrollo docente es esencial considerando que la enseñanza es exigente y que las expectativas del profesorado pueden evolucionar con el tiempo (OCDE, 2023). Ofrecer oportunidades e incentivos para el desarrollo profesional (DP) permite que el profesorado se actualice, desarrolle y amplíe su conocimiento y comprensión de la enseñanza, y mejore sus habilidades y prácticas.

El DP entendido como el proceso mediante el cual un docente adquiere o perfecciona las habilidades, conocimientos y/o actitudes para mejorar la práctica educativa (Mitchell, 2013), supone abarcar toda una gama de actividades (cursos formales, seminarios, conferencias, talleres, capacitación en línea, tutorías, supervisión). Este puede impartirse dentro de las propias instituciones educativas o con la colaboración de institutos de formación e instituciones de educación superior, y se financia por gobiernos, empleadores o a título personal, o en acuerdos de cofinanciación. Se trata de una formación en donde se combinan enfoques y metodologías pedagógicas exitosas para alcanzar los objetivos propuestos, lo que exige un equilibrio entre las demandas de los docentes y las oportunidades que ofrecen los centros educativos (Ibrahim et al., 2020). Sin embargo, los beneficios del DP dependen de la calidad de los programas y de la retroalimentación y del apoyo que se asigne para su seguimiento (OCDE, 2023).

Los responsables de las políticas educativas son conscientes de la necesidad de invertir en recursos físicos y humanos para una formación de

calidad. Por ejemplo, en Estados Unidos se destina el 7.6% del presupuesto de los distritos escolares, en Canadá el 6.5% de los días laborales para actividades de DP (Basma y Savage, 2018); en la Unión Europea, más de la mitad de los fondos se canaliza a la investigación e innovación, aunque advierten sobre la necesidad de continuar mejorando la formación del profesorado y el atractivo de la profesión docente (Eurydice, 2017); y en América Latina, la escasa relevancia y calidad de la oferta formativa, la desvinculación con la realidad de las escuelas (Vaillant, 2016) y de la propia realidad socio-histórica de la región (Bennasar, 2020), hacen que el presupuesto se diluya en una formación de bajo impacto.

Son los gobiernos de los países quienes establecen el marco general que da forma a sus sistemas educativos y definen su funcionamiento, por lo que, sus esfuerzos tienen que encaminarse para conseguir que las escuelas sean lo más efectivas posibles. Es el profesorado uno de los pilares fundamentales para alcanzar una educación de calidad. La mayoría de las evaluaciones sobre los sistemas de enseñanza y educación colocan al docente en el centro de cualquier intento de producir un cambio positivo en los resultados de aprendizaje de los estudiantes (Hattie, 2015; Pamies et al., 2022, Santos, 2020).

Una adecuada formación hace que el profesorado en ejercicio desarrolle y mejore sus competencias docentes para responder a los problemas derivados de su práctica diaria en el centro educativo (Cabero-Almenara y Palacios-Rodríguez, 2020). De hecho, los programas de formación de DP más efectivos se caracterizan por ser concretos, abordan situaciones reales, permiten a los docentes que elijan sus actividades, fomentan el trabajo colaborativo, brindan tutorías de seguimiento y reciben el apoyo de un liderazgo escolar eficaz (Aldahmash et al., 2019).

De esta manera, "la naturaleza del acto pedagógico en el contexto de la Educación, comporta una complejidad relacional sistémica que compromete la formación y la identidad profesional del docente" (Rojas y Martínez, 2023, p.249). Es en esta complejidad relacional donde la formación es multifacética e influye en el conocimiento, creencias y acciones de los docentes, pero a la vez, en el aprendizaje de los estudiantes. En consecuencia, ha de primar una formación continua que vele por la realización y cumplimiento del DP "que permita la convivencia en una sociedad democrática, justa y equitativa" (Van Grieken, 2016, p.32). También es en ese continuo en donde la formación integral pasa por lo pedagógico, ético, científico, tecnológico y humanístico (Suárez et al.,

Recibido: 26/03/2023 Aceptado: 03/05/2024

2019), a lo que Grau et al. (2009) añaden: "es posible alcanzar la excelencia educativa, si existe una formación integral del profesorado" (p.84).

El Estudio Internacional de Enseñanza y Aprendizaje (TALIS, *Teaching and Learning International Survey*, en inglés) que da voz a maestros y directores a aspectos relacionados con la formación docente que han recibido, y a sus creencias y prácticas educativas (OCDE, 2021), desvela que, los docentes de Educación Primaria consideran que la formación y el apoyo son excelentes herramientas para fomentar la confianza de los profesores. En cambio, los maestros de Educación Secundaria y Bachillerato son quienes menos formación reciben durante sus años en ejercicio a pesar de contar con una larga experiencia laboral fuera del ámbito educativo (OCDE, 2021) siendo contados también los programas dirigidos a docentes noveles. En España, ya en el 2015, el XXI Encuentro de Consejos Escolares Autonómicos y del Estado dedicó su monográfico al profesorado del siglo XXI afirmando que "la calidad de un sistema educativo no puede exceder la calidad de sus profesores" (Ministerio de Educación, Cultura y Deporte, 2015, p.12).

La necesidad de sintetizar la formación recibida para el DP de docentes en ejercicio es el interés del presente estudio porque "la docencia se convierte en una de las profesiones más dinámicas en cuanto a necesidades de transformación individual y colectiva, pero, paradójicamente, en una de las más conservadoras para aceptar y comprender las nuevas demandas laborales y sociales" (Alves, 2003, p.7).

Es por ello que el objetivo general del estudio es realizar una revisión sistemática (RS) de artículos científicos para analizar la formación que reciben docentes en ejercicio de educación primaria, secundaria y bachillerato para su DP, qué caracteriza a esos modelos de formación y cuál ha sido su impacto.

La investigación se justifica porque una búsqueda de RS previas en DP apuntan a la universidad con relación a las redes sociales como espacio para que los profesores se involucren en el aprendizaje profesional (Luo et al., 2020) o si el trabajo entre colegas universitarios conduce a acciones reales para construir un proceso de mejora continua y cómo esto influye en la esfera individual y organizativa de las universidades (Gast et al., 2017). Otras revisiones más recientes estudian el DP de profesores universitarios que imparten STEM (Surahman y Wang, 2023) o los efectos potenciales del DP docente en el rendimiento de los estudiantes de inglés como lengua extranjera (Zeng, 2023) o en profesionales de la educación

estadounidenses sobre desigualdad y racismo (Matschiner, 2023). Para Almazroa y Alotaibi (2023) es necesario en este tipo de revisión un mayor número de bases de datos a consultar y un análisis que incluya la formación específica para la enseñanza de habilidades del siglo XXI como es la creatividad.

Asimismo, las contadas RS focalizadas en las escuelas abordan temas concretos sobre el DP y las TIC para una enseñanza a distancia (Ibrahim et al., 2020), la evaluación de la coherencia de los programas de DP y su eficacia (Lindvall y Ryve, 2019), o el video para el DP de docentes en ejercicio de Educación Especial (Morín et al., 2019). Solo la revisión de Bascopé et al. (2019) se aproxima a comprender el DP desde un enfoque más general, pero se limita a estudiar a los docentes de infantil.

El artículo aporta una nueva RS centrada en las escuelas porque la mayoría de la RS sobre DP se llevan a cabo en la universidad, o las que sí lo hacen, ofrecen una información muy específica sobre la formación impartida o solo se centran en docentes de infantil. Además, se diferencia de otras revisiones porque muestra una visión general de la formación ampliando las bases de datos consultadas sin restringir los hallazgos a un país en específico, estudia en qué se forman los docentes y cómo impacta en ellos dicha formación, propone un análisis diferenciado por etapas educativas así como un análisis cualitativo de los estudios primarios a partir de categorías emergentes.

La RS pretende contribuir con un mapeo sobre la formación para el DP de docentes en ejercicio de escuelas y sus implicaciones educativas para una toma de decisiones basada en la evidencia.

Método

Una revisión sistemática (RS) utiliza métodos repetibles para encontrar, seleccionar y sintetizar toda la evidencia disponible. Responde a una pregunta de investigación claramente formulada y establece explícitamente los métodos utilizados para llegar a la respuesta (Nunn y Chang, 2020). La presente RS tiene un carácter exploratorio debido a la necesidad de hacer un mapeo general sobre el tema de interés con la finalidad de localizar brechas en la literatura (Fernández-Sánchez et al., 2020). Para llevar a cabo la revisión, se han cumplido con los criterios y fases de la declaración PRISMA (Newman y Gough, 2020; Page et al., 2021).

Recibido: 26/03/2023 Aceptado: 03/05/2024

Las preguntas de investigación que han guiado la RS cualitativa han atendido el modelo PICO, específicamente en su variante PO (Booth et al., 2019):

- PI1 ¿En qué se forman los docentes de Educación Primaria, Educación Secundaria y Bachillerato para su DP?
- PI2 ¿Cuáles son los modelos o enfoques de formación y en qué consisten?
- PI3 ¿Cuál ha sido el impacto de la formación recibida? (Tabla I).

La revisión de la literatura se llevó a cabo en cuatro bases de datos: dos generales como son SCOPUS y *Web of Science* (WoS), otra específica como es *Education Resources Information Center* (ERIC), y Dialnet, por tratarse de un portal que recopila y proporciona acceso fundamentalmente a documentos de literatura científica hispana. Según Bramer et al. (2017) lo recomendable es la combinación de bases de datos generalistas porque es en ellas donde se obtiene el mayor porcentaje de recuperación de los artículos.

Los descriptores y la cadena de búsqueda empleados han sido *professional development* OR *training* OR *teachers* combinados con *primary education* OR *secondary education* OR *high school* OR *school* en el título, palabras clave o resumen según los requerimientos de cada base de datos. En cambio, para los términos *university*, *college y early childhood education* se empleó el operador boleano NOT. Estos descriptores se seleccionaron del Tesauro Europeo de Educación, ERIC y de otras RS previas relacionadas con el tema (Gast et al., 2017; Ibrahim et al., 2020; Luo et al., 2020).

TABLA I. Formulación preguntas que guían la RS

Variante PO	PI1	PI2	PI3
(P) *Population/ Phenomena*	Formación docente de Educación Primaria, Secundaria y Bachillerato	Formación docente de Educación Primaria, Secundaria y Bachillerato	Formación docente de Educación Primaria, Secundaria y Bachillerato
(O) *Outcome*	Objetivos y contextos de formación	Modelos o enfoques de formación	Impacto

Fuente: elaboración propia. Nota: PI1= pregunta de investigación 1, PI2= pregunta de investigación 2, PI3= pregunta de investigación 3.

Los criterios de inclusión han sido estudios primarios: (1) publicados entre 2010-2021; (2) empíricos; (3) escritos en inglés o en español; (4) a texto completo; (5) artículos de investigación; (6) con revisión por pares; (7) de docentes en ejercicio (8) de primaria, secundaria o bachillerato (9) quienes han recibido formación para su DP según la definición de Mitchell (2013). Por su parte, los criterios de exclusión han sido: (1) estudios teóricos y otras RS; (2) centrados en formación para estudiantes de colegios; institutos o universitarios; (3) formación en otras disciplinas diferentes a Educación; (4) formación sólo para directivos y personal administrativo; (5) formación para futuros docentes (formación inicial); (6) profesores de infantil y universitarios; (7) estudios que abarquen el tramo no universitario en donde la información relativa a educación primaria, secundaria y bachillerato no es clara o suficiente o (8) donde el diseño de metodologías, técnicas o recursos no se acompañen con formación. Se decidió excluir a docentes de infantil, ya que existe una RS reciente sobre los enfoques pedagógicos más apropiados para el desarrollo sostenible en etapas tempranas (Bascopé et al., 2019).

La evaluación de la calidad metodológica de los estudios primarios siguió las recomendaciones de Alexandre (2020), para quien lo importante es que los investigadores puedan estar relativamente de acuerdo en la gravedad de una amenaza a la calidad del diseño de un estudio. Por tanto, solo se incluyeron artículos de revistas con un proceso de revisión por pares.

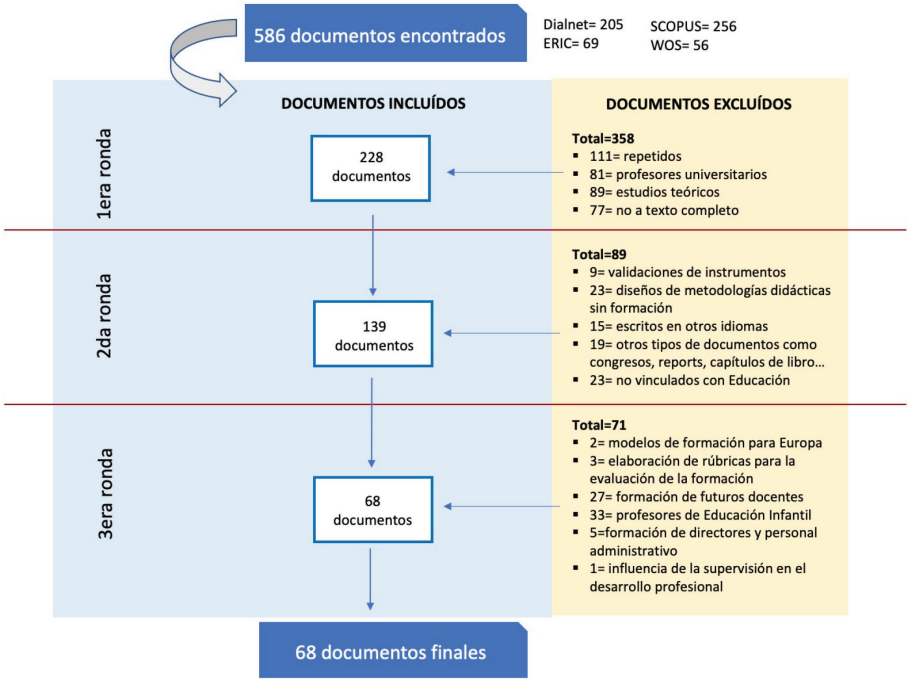
Durante el proceso de revisión y sus fases, cada investigador codificó independientemente una base de datos para luego revisar y verificar la información de otro compañero. El consenso en las diferentes rondas pasó de un acuerdo inter-observadores de un 86.5% hasta la totalidad resolviéndose cualquier desacuerdo a través de un proceso reflexivo constante. La figura I representa el diagrama de flujo según PRISMA (Page et al., 2021). De un total de 586 documentos encontrados solo 68 cumplieron con los criterios de inclusión.

El análisis de los 68 documentos finales se realizó a través de un manual de codificación conformado por *variables extrínsecas* (año de publicación y procedencia de los autores) *variables metodológicas* (muestra y método de investigación empleado) y *variables sustantivas* (en qué se forman los docentes, características de los modelos de formación recibida e impacto) (manual de codificación.xlsx).

Las variables extrínsecas y metodológicas se analizaron con estadística descriptiva (frecuencias y porcentajes) y con ayuda de Excel pero el análisis de las variables sustantivas fue temático (Morgan, 2022) siguiendo

Recibido: 26/03/2023 Aceptado: 03/05/2024

FIGURA I. Diagrama de flujo



Fuente: elaboración propia.

la propuesta de Flick (2018) que implica identificar patrones de ideas, conceptos o temas similares para luego establecer relaciones e integrar la información entre sí con los fundamentos teóricos del tema. Todas las categorías fueron emergentes porque favorecen una mayor comprensión del objeto de interés (Cypress, 2017) cumpliendo el criterio de análisis inductivo de abajo hacia arriba (*bottom-up)* (Bingham y Witkowsky, 2022). La triangulación entre los cuatro investigadores del estudio (Flick, 2018) permitió ajustar dichas categorías: dos investigadores realizaron la primera codificación de manera independiente, el tercero se encargó de poner en común ambas codificaciones (similitudes y diferencias) y el cuarto repitió el procedimiento intentando reducir las diferencias hasta alcanzar un acuerdo total con los otros tres compañeros. Estos roles se intercambiaron para la categorización de las tres preguntas de investigación que guiaron la RS.

Resultados

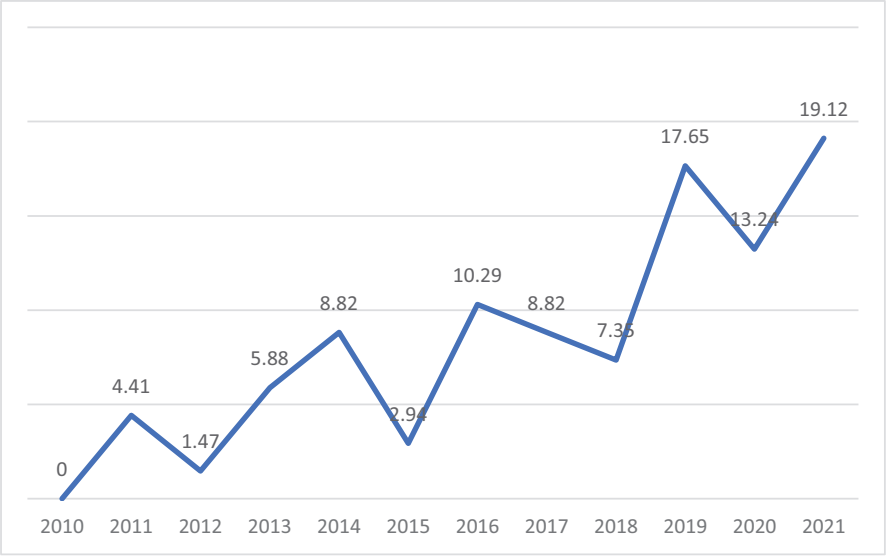
A continuación, se presenta una descripción general de las características de los artículos seleccionados, en qué se forman los docentes para su DP, bajo qué modelos y enfoques se imparte la formación y cuál ha sido su impacto. Para facilitar la lectura, los 68 estudios finales se han identificado con un número y estos pueden consultarse en el Anexo I.pdf.

Características generales de los artículos

El 50% de los estudios analizados se concentra entre 2019-2021 y el 76.47% desde el año 2016. Entre 2010 y 2021 se observa un incremento interesante en la formación de los docentes siendo el año 2021 con la mayor producción científica (19.12%) (Figura II).

Europa y América aglutinan el 65.71% de estudios relacionados con la formación para el DP de docentes en ejercicio de escuelas (Figura III).

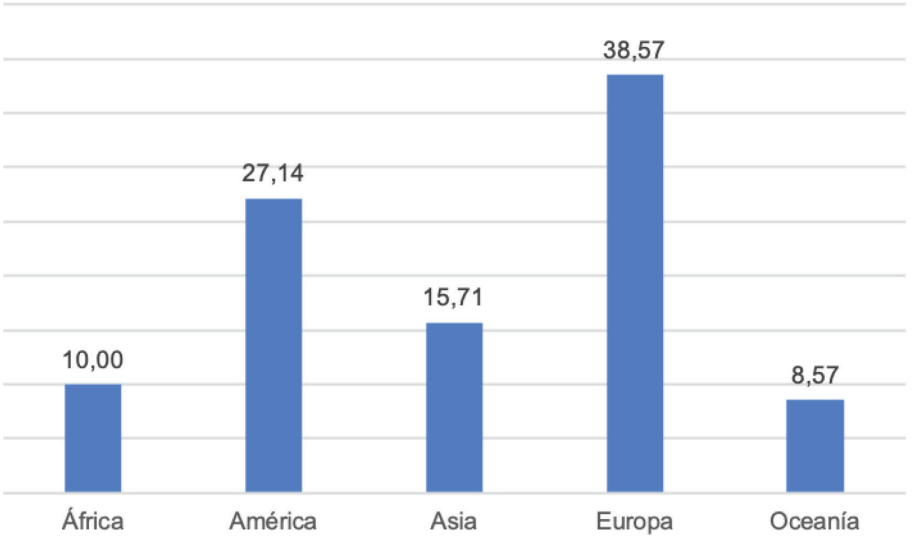
FIGURA II. Producción de artículos seleccionados en función del año



Fuente: elaboración propia. Nota: figura que representa el porcentaje de producción de artículos en función del año.

Recibido: 26/03/2023 Aceptado: 03/05/2024

FIGURA III. País de procedencia de los artículos seleccionados



Fuente: elaboración propia.
Nota: figura que representa el porcentaje de publicaciones por continente.

España es el país con mayor número de publicaciones con un 14.29% (26, 33, 36, 38, 41, 47, 52, 54, 55, 56) seguido de cerca por Estados Unidos con un 10% (7, 25, 28, 43, 67, 59, 61) y luego por Australia (5, 20, 27, 48, 68, 57) y Sudáfrica (1, 4, 11, 12, 13, 66) con apenas un 8.57%. Otros países con menor participación son Reino Unido (9, 19, 21, 23) y Chile (42, 46, 50, 51) con un 5.71%; Países Bajos (16, 18, 24), Chipre (29, 64, 58) y Ecuador (34, 39, 44) con un 4.29%; y Jordania (8, 53), Colombia (30, 37) e Israel (2, 3) con 2.86%.

No sólo los docentes han recibido formación sino también directores (8, 10, 14, 66) y personal administrativo (25) además de maestros noveles (2), mentores (23) y auxiliares (*teachers assistants*) (19). En algunos estudios participaron hasta 410 profesores en una formación sobre aprendizaje basado en el juego (64) o 616 para examinar los beneficios de la formación continua del profesorado (65).

Por otra parte, la mayoría de los artículos se han llevado a cabo en Educación Primaria (43.42%) y en Educación Secundaria Obligatoria (30.26%) no siendo tan frecuentes en Bachillerato (2.63%) con apenas sólo dos estudios (2,30). Un 23.68% indica que la formación se ha

realizado en las escuelas sin ofrecer resultados específicos por etapas educativas.

Para evaluar la formación, prevalece la investigación cualitativa (36.76%) aunque un 76% de estos estudios no especifica el método empleado (2, 18, 26, 30, 31, 36, 45, 48, 50, 51, 55, 57, 61, 63, 64, 68, 65, 66, 67). El estudio de caso es el que más se ha utilizado (5.88%) (6, 20, 22, 28, 29, 58) combinado con otros como *Design-based research* (29) y la etnografía (58). En cambio, en investigación cuantitativa (29.41%), el 55% de los estudios se ha llevado a cabo con diseños exploratorios-descriptivos (4, 37, 38, 40, 42, 44, 47, 53, 54, 60, 62), un 20% son correlacionales (33, 34, 41, 49) y un 10% refieren a análisis factorial (32, 52) y clúster (35, 46), respectivamente. Sólo siete estudios han empleado métodos mixtos (10.29%) (8, 17, 19, 21, 39, 43, 56).

¿En qué se forman los docentes para su DP?

La formación en las escuelas gira en torno a cuatro grandes categorías. La que más prevalece es la formación no curricular (45.21%) en donde los docentes se forman en conocimientos, habilidades y competencias no vinculadas con las asignaturas que imparten; seguida por modelos de formación colaborativos (21.92%) que se sustentan en la ayuda mutua entre profesores. De igual forma, pero con menor frecuencia, un grupo de estudios hace un mapeo de la formación recibida y las necesidades de formación y expectativas (16.44%) y otro grupo reporta una formación específica dirigida a las asignaturas que los maestros transmiten en sus aulas (16.44%) (Tabla II).

¿Cuáles son los modelos o enfoques de formación y en qué consisten?

Los modelos de formación en cascada tienen efectos positivos en las actitudes y prácticas de maestros de secundaria cuando se acompañan de un seguimiento y de talleres periódicos que ofrecen un contacto continuo entre formadores y maestros. De esta manera, toda la escuela participa activamente del DP. En este tipo de formación, el apoyo de los directores es vital para la mejora profesional del profesorado (10) siendo igualmente interesante cómo desempeñan esta gestión (11).

Recibido: 26/03/2023 Aceptado: 03/05/2024

TABLA II. Objetivos y contextos de la formación docente

Objetivos y contextos de la formación docente: categorías y códigos emergentes	
1. Mapeo: formación recibida, necesidades formativas y expectativas (16.44%)	
Indagar qué formación están recibiendo los docentes	- Orientación para políticas de enseñanza-aprendizaje (1) - Impacto días capacitación en el DP (21)
Detección de necesidades de formación (formación permanente, cursos de formación al profesorado importancia formación continua)	- Formación en inglés para zonas rurales (7) - Papel directores en la gestión del DP (11) - Necesidades formación permanente (53) - Importancia formación continua (65,66)
DP docente (eventos culturales, sociales, entorno…)	32,33
Creencias sobre la docencia en el desempeño docente	32,35
2. Formación no curricular (45.21%)	
Mentoring, tutoría	- *Mentoring* (2,61) - Creación de comunidades de co-aprendizaje compartido (67)
TIC, Competencia digital (informatización y alfabetización informacional)	- Desarrollo competencias digitales (6) - Usos teléfonos móviles (9) - TIC (12,34,43) - Uso videoconferencias sincrónicas (20) - Creación de redes sociales para el DP (68)
Enseñanza y aprendizaje	- Actitudes hacia la enseñanza (10) - Proceso de enseñanza-aprendizaje (17) - Actitudes docentes auxiliares (19)
Investigación aplicada	18
Liderazgo docente y coaching	31,42
Innovación docente	36,38,44,45,47
Resolución de conflictos	37
Comunicación y aprendizaje basado en el juego (gamificación)	64
Labor o acción tutorial	54
Motivación escolar	46
Inclusión educativa y educación intercultural	- Educación intercultural (15) - Inclusión educativa y la atención a la diversidad (23,49,51,55,62,63)
Evaluación (técnicas para mejorar la evaluación en los estudiantes)	56
3. Formación curricular (16.44%)	
Orientación para la vida (justicia social, derechos humanos, conciencia del medio ambiente…)	13

(Continuada)

Recibido: 26/03/2023 Aceptado: 03/05/2024

TABLA II. Objetivos y contextos de la formación docente *(Continuada)*

Conocimiento didáctico (área de Naturaleza, Ciencia y Tecnología)	40
Matemáticas	3,52
Estadística	30
Economía	4
Inglés	5,39,50
Química	22
Actividad Física o deportiva	41
TIC	8
4. Modelos de formación colaborativos (21.92%)	
Co-design	27,29
Co-teacher	28,57,58,59,60
Peer to peer	26
Peer-observation	24,25
Colaboración escuela-universidad	14
Validación del modelo instruccional 4C/ID	Centrado en tareas de aprendizaje, información de apoyo, información de procedimiento y práctica de tareas parciales (16)

Fuente: elaboración propia.

Cuando la formación para el DP se sustenta en andragogía y favorece la participación activa y colectiva del profesorado (17), los resultados son igualmente prometedores. Los hallazgos evidencian que este modelo de formación mejora la planificación, los métodos de enseñanza, la gestión del aula y la cooperación entre docentes porque modifica sus prácticas y estilos de enseñanza (17).

De igual forma, son efectivos los programas para la enseñanza de las matemáticas con un enfoque orientado en el desarrollo de competencias a lo largo de la vida (16) o también los resultados son igualmente interesantes con modelos colaborativos (Tabla III), pero no siempre es así: si el modelo de formación es impuesto por las administraciones educativas, la calidad tiende a decaer (13,21).

Por último, una de las principales condiciones que incide de manera positiva en la eficacia de los modelos de formación, es el número de cursos recibidos por el profesorado. Por ejemplo, quienes han asistido a dos o más cursos universitarios se perciben más efectivos en la aplicación

Recibido: 26/03/2023 Aceptado: 03/05/2024

TABLA III. Tipos de modelo de formación colaborativos

Tipos modelos-colaborativos	Resultados de la formación
Co-diseño *(co-design)*	- Mayor conciencia zona de desarrollo próximo (27) - Necesidad de incrementar medidas cuantitativas de autoeficacia y autoinformes antes y después de la formación (27) - Incidencia positiva en la cultura escolar (27) - Promueve la reflexión pedagógica (90.3%) (29) - Solo 35.5% de la experiencia se difundió en congresos (29)
Co-enseñanza *(co-teaching)*	- Impacto positivo en la docencia (58) - Invita a la programación conjunta (57,60) - A mayor edad mayor percepción positiva del *co-teaching* (61) - Favorece un modelo inclusivo para estudiantes con necesidades educativas especiales (55, 57, 58, 59, 61) - Fortalecimiento tutores experimentados y noveles (67) - Incrementa la sensación de comunidad (31) y la relación entre docentes (comunicación, colaboración, confianza, relación personal, cercanía) (31, 57, 59, 60) - Rompe con el aislamiento del profesorado (31) - Potencia el liderazgo docente (31) - Se focaliza en el aprendizaje y no la enseñanza (28)
Formación entre pares (*peer to peer)*	- Garantiza la implementación del programa y su sostenibilidad en el tiempo (26)
Observación entre pares *(peer-observation)*	- Promueve el trabajo en equipo y la reflexión de la práctica (24) - Mejora la comprensión y la aplicación de estrategias de enseñanza (24) - Permite anticipar las reacciones de los estudiantes durante las clases (24) - Fortalece la confianza (24) - Dificultades para elegir a qué colegas observar debido a la incompatibilidad de horarios (25)
Formación escuela-universidad	- Promueve el diseño de programas en mentoría e investigación (14) - Condiciones que inhiben o promueven una formación colaborativa-equitativa: (a)generar compromiso interno entre los diferentes actores involucrados, (b) implicación de coordinadores y directores y (c)comunidad centrada en prácticas basadas en la indagación (14)

Fuente: elaboración propia.

de métodos y estrategias para la enseñanza del inglés como segunda lengua en comparación con aquellos que tienen menos formación (7). Asimismo, los directores de las escuelas son otra condición clave para ayudar al crecimiento profesional de sus profesores. En la medida en que participan en los programas de formación se observan cambios positivos en el desempeño de sus docentes (66). Además, los cursos más efectivos son aquellos que satisfacen directamente las necesidades individuales y de la escuela (21) y que se preocupan por atender a los

docentes más experimentados (1). Sin embargo, la calidad disminuye cuando la formación presenta deficiencias e incoherencias con relación a la planificación, contenidos, financiación, recursos (relación calidad-cantidad), seguimiento, aplicabilidad (13), y temporalidad (ej. impartición fuera del calendario laboral) (21).

¿Cuál ha sido el impacto de la formación recibida?

El impacto de la formación recibida incide en seis categorías emergentes que refieren a la enseñanza de las asignaturas del currículum, la competencia digital, el bienestar socioemocional, los docentes noveles y auxiliares, la comunidad educativa y la identidad profesional.

Categoría 1. Impacto de la formación en la enseñanza. En general, la formación incide en un mayor conocimiento teórico y didáctico sobre las asignaturas que se imparten, a favor de un cambio actitudinal positivo con respecto al proceso de enseñanza-aprendizaje. También se señala que es necesario una formación especializada en matemáticas para estudiantes de altas capacidades, y que hay asignaturas que, por su complejidad, como es el caso de química, requieren de una formación más prolongada en el tiempo (Tabla IV).

Categoría 2. Impacto de la formación en la competencia digital. La formación en competencias digitales se ha disparado entre el profesorado como consecuencia de la pandemia COVID-19 (34), en especial en nociones de alfabetización informacional, elaboración de contenidos digitales y seguridad en Internet (34).

En líneas generales, la formación de los docentes en TIC apunta a resultados positivos y negativos, pero también a necesidades específicas.

Por una parte, la competencia digital mejora el conocimiento y las habilidades de los profesores, especialmente en el compromiso profesional, empleo de recursos, planificación de actividades y mejora de las evaluaciones de los estudiantes (6). Otros resultados positivos ponen en evidencia el reconocimiento del potencial educativo de los móviles en la enseñanza y el aprendizaje de las ciencias (9). De la misma manera, los docentes han demostrado ser competentes al combinar habilidades técnicas, pedagógicas e interpersonales para impartir clases online: comunicar a través de las pantallas, establecer relaciones con los estudiantes mediadas por la tecnología y diseñar clases que fomenten la participación de los alumnos (20).

Recibido: 26/03/2023 Aceptado: 03/05/2024

TABLA IV. Resultados en la enseñanza de distintas materias/asignaturas

Materias/ asignaturas	Resultados de la formación
Ciencia, naturaleza y tecnología	- Fortalecimiento del pensamiento científico y tecnológico mediante prácticas de indagación (40) - Huertos escolares: (a)alto nivel de motivación, (b)conciencia de la importancia de la participación activa por parte de los estudiantes, (c)necesidad de formación en aspectos técnicos (ej. técnicas agrícolas, cultivos, semillas, suelos, etc.) y didácticos (ej. uso escolar, actividades, su planificación y evaluación), y (d)integración en el currículo de Ciencias de la Naturaleza en todas las áreas (47)
Matemáticas y estadística	- Escasa formación en matemáticas para la enseñanza a estudiantes con altas capacidades (3) - Mejora la actitud de las matemáticas cuando se incluyen actividades/estrategias reflexivas y activas (ej. papel de los problemas matemáticos en el aula o función de las preguntas y los diálogos en el aprendizaje) (52) - Mejora la actitud y la enseñanza de la estadística cuando se reflexiona sobre la propia vivencia como estudiante (30)
Economía	- Mejora de la actitud hacia la enseñanza de la economía (4) - Mayor aprendizaje en contenidos relacionados con problemas económicos contemporáneos (4)
Inglés	- Incremento del aprendizaje cuando existe una coordinación eficaz entre profesores, administradores y directivos (5)
Lengua	- Aprendizaje del potencial de las narrativas y mayores niveles de conciencia meta-semiótica (50) - Poner en valor la literatura en una era multimodal (50)
Química	- Un año de formación no es suficiente para el desarrollo de docentes líderes competentes (22)

Fuente: elaboración propia.

Como contraparte, no siempre la formación ha logrado que el profesorado sea competente en lo digital. Por ejemplo, la formación en el uso del mini-Pat no respondió a las necesidades de los docentes ni tampoco amplió el repertorio de conocimientos y habilidades en el uso de las TIC (12). Pese a que los programas de DP ayudaron a los maestros a mejorar sus habilidades y conocimientos sobre la tecnología, algunos factores incidieron de forma negativa en su eficacia, entre ellos: el momento y el modelo de capacitación (ej. al finalizar la jornada, cursos impuestos por las administraciones), la falta de motivación y seguimiento, una cultura escolar que no reconoce el DP y la sobrecarga de trabajo de los docentes (8).

Con respecto a las necesidades detectadas, el profesorado requiere tiempo para apropiarse del conocimiento y así mejorar su práctica docente digital (68). Las TIC siguen siendo un motivo de diferencias entre el claustro de profesores de un mismo centro educativo y en la dotación de espacios tecnológicos entre las distintas escuelas (44). Aunque se generen diseños de entornos de aprendizaje (38), se valore favorablemente la figura del coordinador TIC (43) y el porcentaje de formación en tecnología sea elevado (ej. pizarras digitales) (36), parece que, cuando el apoyo de las escuelas no es suficiente o adecuado (43), genera en los docentes frustración (38), lo que incide en propuestas didácticas de baja calidad (36).

Por último, otras variables a tener en cuenta que pueden influir a favor o en contra para el desarrollo de la competencia digital es la actitud y el nivel de alfabetización mientras que el estado civil, nivel educativo, especialidad y cargo profesional no inciden así como tampoco los modelos de formación (62).

Categoría 3. Impacto de la formación en el bienestar socio-emocional. Se observa en la Tabla V cómo recibir formación para el autocuidado físico y el desarrollo de competencias socioemocionales, repercute también en el bienestar de estudiantes y familias. Los docentes igualmente demandan formación en educación inclusiva debido a sus implicaciones positivas en toda la comunidad educativa.

Categoría 4. Impacto de la formación en docentes noveles y auxiliares. Aunque son contados los estudios que hacen alusión a este impacto, los docentes noveles inician su DP con formación en la función tutorial por ser un factor de calidad para el acompañamiento socio-afectivo y cognitivo de sus estudiantes. En ella, son conscientes de la propia naturaleza y complejidad de la tutoría, la conexión esencial entre teoría-práctica, y la importancia de desarrollar habilidades reflexivas que potencien un pensamiento crítico en sus alumnos (2). En cambio, el mayor impacto de la formación en docentes auxiliares (*teacher assistant*) es en lo personal, pero parece tener poco o ningún impacto en el desarrollo de sus carreras, remuneración y reconocimiento laboral (19).

Categoría 5. Impacto de la formación en la comunidad educativa. Cuando los maestros se forman en investigación, desarrollan una mayor conciencia de la escuela como una organización compleja, del papel que juega el personal administrativo y del potencial de la indagación y sus beneficios para la innovación en el aula (18). En esta misma línea, la formación en liderazgo dirigida a directivos para el manejo del tiempo y

Recibido: 26/03/2023 Aceptado: 03/05/2024

TABLA V. Resultados en bienestar socioemocional

Formación en	Resultados de la formación
Bienestar físico	- A mayor bienestar físico mayor motivación hacia la práctica (41)
Manejo del fracaso y la frustración	- Aprendizaje de metodologías emocionales centradas en la vivencia (38)
Motivación escolar y familia	- Es una responsabilidad compartida con la familia y no un proceso continuo que recae en el propio docente (46)
Resolución de conflictos o violencia escolar	- Identificación de factores para el manejo de conflictos en el aula: autoconocimiento, autorregulación y motivación (37) - Identificación de factores que determinan el éxito escolar y relaciones sociales no violentas: autocontrol y motivación (37)
Educación inclusiva	- Falta de formación durante la formación inicial (51) y con un enfoque humanista (23) - Ausencia de formación para la transición de la enseñanza del alumnado indígena (48) - Necesidad de fortalecimiento del vínculo escuela-familia (49, 63) - Identificación de estrategias de comunicación para aumentar la participación de los padres de escuelas rurales (7)

Fuente: elaboración propia

de los conflictos, evidenció una mejora de la gestión del centro educativo (42). Sin embargo, la formación en escuelas rurales para la enseñanza del inglés como segunda lengu, no logró alcanzar sus objetivos debido a las dificultades de comunicación entre docentes y familia, situación que impidió una participación más activa (7).

Categoría 6. Impacto en formación de la identidad docente. Cuando los docentes pueden elegir cómo mejorar su DP, la mayoría se forma en su campo o especialidad (25.7%) o en TIC (10%) o en cuestiones disciplinarias (4.8%) (32) y gestión del aula (4.8%) (32,35). Otro hallazgo importante es que el DP refuerza la identidad profesional del profesorado en donde influyen los rasgos de personalidad (35), las creencias con respecto al propio desempeño, los años de experiencia laboral y de especialidad (32). Por ejemplo, profesores con varios años de experiencia en docencia se sienten mejor preparados para el desempeño de sus funciones educativas (ej. tareas relacionadas con la planificación y la docencia) (32), lo que fortalece la dimensión personal y profesional propias del DP. Sin embargo, una valoración negativa de la profesión (33) hace que la efectividad de los programas de DP disminuya.

Discusión y conclusiones

El DP es una actividad educativa que ayuda a mantener, desarrollar o mejorar el conocimiento, la resolución de problemas y la experiencia para lograr una enseñanza de calidad. Por este motivo en el presente estudio se llevó a cabo una RS de artículos científicos para analizar la formación que reciben docentes en ejercicio de educación primaria, secundaria y bachillerato para su DP, qué caracteriza a esos modelos de formación y cuál ha sido su impacto.

En líneas generales, la formación en las escuelas se ha incrementado con más del 50% de los documentos analizados entre el 2019 y el 2021, dato que confirma la demanda de programas de formación durante la pandemia COVID-19 y que es recogido en informes de la OCDE (2021) y en investigaciones de distintos países (Irby et al., 2022; Lo y To, 2023; Perry, 2023; Wu et al., 2022).

Son los docentes de Educación Primaria quienes reciben mayor formación (43.42%) seguidos por Educación Secundaria (30.26%) y con solo un 2.63% de profesores de Bachillerato, situación que es similar en los países de la OCDE (2021). En España, el 48% del profesorado ha recibido formación sobre el contenido de las materias, pedagogía y prácticas en aula frente al 79% de estos países (OCDE, 2021).

Con respecto a la pregunta ¿En qué se forman los docentes para su DP? (PI1), predomina el desarrollo de competencias digitales (6, 9, 12, 20, 34, 43, 68), hallazgo que también comparten otros autores (Cabero-Almenara y Palacios-Rodríguez, 2020) y que se ha incrementado aún más por la COVID-19 (Perry, 2023). La necesidad de formación en el ámbito de la atención a la diversidad (15, 23, 49, 51, 55, 63), coincide con el reporte de la OCDE (2021), que afirma que existe un desafío con relación a la construcción de aulas más inclusivas y de docentes con más experiencia (Wächter y Gorges, 2023). Por último, se evidencia un creciente interés en modelos de formación colaborativos (14, 16, 24, 25, 26, 27, 28, 29, 57, 58, 59, 60) porque favorecen la retroalimentación, el desarrollo de buenas prácticas (Colson et al., 2021) y entornos compartidos de reflexión (Westbroek et al., 2019). Para Pedroza y Reyes (2019) "… [estos modelos] conducen a una autoevaluación de la formación continua, lo que impacta en el éxito de los aprendizajes" (p.426).

Por otra parte, ¿Cuáles son los modelos o enfoques de formación? (PI2). Es interesante cómo los modelos de formación colaborativos son objeto

Recibido: 26/03/2023 Aceptado: 03/05/2024

de estudio en sí mismos, a la vez que son un tipo de formación que se trabaja en las escuelas (14, 24, 25, 26, 27, 28, 29, 31, 55, 57, 58, 59, 60, 61, 67) y que se van asentando cada vez más como una línea de investigación prometedora (Colson et al., 2021). Asimismo, los hallazgos encontrados apoyan investigaciones recientes que también confirman la eficacia de los modelos de formación de cascada (Savard y Cyr, 2018) y a lo largo de la vida (Reychav et al., 2023), aunque ambos enfoques no están exentos de la dificultad que supone formar a un número importante de docentes.

Con relación a ¿Cuál ha sido el impacto de la formación reciba? (PI3), los hallazgos de los estudios analizados indican que el impacto es muy variado, aunque no siempre se valora la formación como suficiente o adecuada. Cuando la formación es eficaz -un 83.83% de los estudios analizados así lo confirma-, primero se produce un cambio en la esfera personal que luego pasa a una fase de aplicación para luego ser evaluado antes de ser adoptado de forma más permanente (Basma y Savage, 2018). En cambio, la insatisfacción del profesorado (16.17%) con la formación recibida se debe a la disconformidad con la planificación (ej. contenido, financiación, recursos, seguimiento) y la incompatibilidad con el calendario laboral y la carga de trabajo (8, 13, 21, 22). A esto se suma la falta de motivación, así como la necesidad de más tiempo para empoderarse y apropiarse de lo aprendido (8, 12, 22, 38, 64, 68). Incluso, estos resultados no favorables se incrementan aún más cuando la formación es en TIC (34, 43, 36, 38).

De igual modo, las necesidades se concentran en una formación académica y técnico-profesional (OCDE, 2021) que permita "afianzar las conexiones entre educación, instrucción e investigación para el desarrollo social" (Suárez et al., 2019, p.19), pero para Barrios y Jiménez (2020), actualmente, la formación está cumpliendo más con una función paliativa que preventiva siendo deseable un nivel más alto de profesionalización. La creación de evaluaciones efectivas que midan las prácticas de los docentes en el aula y que capturen su fortalezas y áreas de mejora son otra prioridad (Poulou et al., 2023).

Implicaciones educativas. Para un DP eficaz se recomiendan oportunidades de formación formales e informales, tiempo para una formación de calidad, impartición de cursos en el horario laboral, fomento de modelos colaborativos entre pares, aumento de programas dirigidos a profesores noveles e inclusión del personal directivo y administrativo de las escuelas.

Respecto a cómo debe organizarse la formación, ha de impartirse en unidades modulares. Los primeros módulos deben ofrecer seguridad a los docentes e incentivar la motivación. Sentadas estas bases, la formación podría sustentarse en investigación-acción-participativa (Pozdeeva, 2018). Se espera que los programas de DP promuevan cambios orientados a lograr una mayor autonomía del profesorado (Choi y Mao, 2021). Esto implica adoptar un enfoque basado en la experiencia con oportunidades para que los maestros aprendan, experimenten e indaguen sobre las prácticas de enseñanza (Choi y Mao, 2021). Es importante que en la formación se trabaje con el análisis de casos prácticos y de estrategias y recursos innovadores (Corujo-Vélez et al., 2021). Una formación de calidad se caracteriza por ser dinámica, interactiva y dialogante "aborda la condición humana como potencialidad" (Nieva y Martínez, 2016, p.14).

Otra implicación educativa es que el profesorado reciba formación por parte de agentes externos a las escuelas o que también observe clases de otros docentes para eliminar barreras didácticas e incorporar nuevas estrategias y procedimientos en sus aulas (Bonnand y Hansen, 2016). Para Rojas y Martínez (2023) "la educación que demanda la sociedad de este tiempo, debe ser pensada desde las universidades para transformar la cultura profesional de los centros escolares" (p.252). Asimismo, resultaría interesante que el profesorado aunara esfuerzos para organizar reuniones interciclos semanales con la finalidad de promover proyectos integrados entre distintas asignaturas.

Limitaciones y prospectiva. Debido a que no se lograron identificar otras RS previas que estudiaran exclusivamente la formación docente para el DP en el tramo no universitario, los hallazgos encontrados se discutieron con investigaciones recientes que abordan el tema de interés. De igual manera, el DP es un constructo complejo que exige de los docentes una formación holística y diversa, pero a la vez especializada en el desarrollo de habilidades, conocimientos y actitudes con la finalidad de mejorar la práctica educativa. Esta complejidad ha estado presente en la variabilidad de enfoques, técnicas y recursos lo que ha supuesto una dificultad para establecer comparaciones mucho más acotadas. Se espera en futuras investigaciones profundizar con un análisis bibliométrico utilizando programas especializados como SciMAT, lo que permitirá identificar la evolución conceptual y las áreas temáticas que concentran la investigación realizada sobre el tema a lo largo del tiempo. Otra línea

Recibido: 26/03/2023 Aceptado: 03/05/2024

podría ser centrarse solo en la formación de los docentes de Bachillerato por ser la menos estudiada o continuar con una revisión de revisiones (*overview*) sobre la formación que reciben los profesionales de la educación en ejercicio para su DP.

Agradecimientos

Esta investigación se ha llevado a cabo en el Grupo de Pedagogía Adaptativa 940424 dentro de proyecto PROFICIENCyIn +EDU, RTI2018-096761-B-I00. MCI/AEI/FEDER, UE, PROFICIENCyIn+E, EDU 2015-63844-R. MINECO/FEDER.

Referencias bibliográficas

Aldahmash, A.H., Alshamrani, S.M., Alshaya, F.S., & Alsarrani, N.A. (2019). Research trends in service science teacher professional development from 2012 to 2016. *International Journal of Instruction, 12*, 163-178. http://doi.org/10.29333/iji.2019.12211a

Alexandre, P.A. (2020). Methodological Guidance Paper: The Art and Science of Quality Systematic Reviews. *Review of Educational Research, 90*(1), 6-23. https://doi.org/ggwsgp

Almazroa, H., & Alotaibi, W. (2023). Teaching 21st Century Skills: Understanding the Depth and Width of the Challenges to Shape Proactive Teacher Education Programmes. *Sustainability*, *15*(9), 7365. https://doi.org/10.3390/su15097365

Alves, E. (2003). La formación permanente del docente en la escuela. El uso universitario de la tecnología para elevar la calidad del docente en el aula. *Investigación y Postgrado, 18*(1), 36-45.

Barrios, C., & Jiménez, B. (2020). Formación dentro y fuera de la escuela. Antítesis o síntesis. *Educar*, *30*, 141-157.

Bascopé, M., Perasso, P., & Reiss, K. (2019). Systematic Review of Education for Sustainable Development at an Early Stage: Cornerstones and Pedagogical Approaches for Teacher Professional Development. *Sustainability*, *11*, 719, 1-16. http://doi.org/10.3390/su11030719

Basma, B., & Savage, R. (2018). Teacher Professional Development and Student Literacy Growth: a Systematic Review and Meta-analysis. *Educ Psychol Rev*, *30*, 457-481. http://doi.org/10.1007/s10648-017-9416-4

Bennasar, M.I. (2020). Los modelos formación docente, un tema en debate permanente. *Revista HumanArtes*, *17*, 1-13.

Bingham, A.J., & Witkowsky, P. (2022). Deductive and inductive approaches to qualitative data analysis. En C. Vanover, P. Mihas y J. Saldaña (Eds.), *Analyzing and interpreting qualitative data: After the interview* (pp. 133-146). SAGE.

Bonnand, S., & Hansen, M.A. (2016). Your device, your classroom: Creating library spaces that support teaching innovation. *C&RL News*, *77*(6), 288-305.

Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., & Shakibazadeh, E. (2019). Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Glob Heal*, *4*(Suppl 1), e001107.

Bramer, W.M., Rethlefsen, M., & Franco, O. (2017). Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic Reviews, 6*(245), 1-12. https://doi.org//10.1186/s13643-017-0644-y

Cabero-Almenara, J., & Palacios-Rodríguez, A. (2020). Formación y competencias del profesorado en la era digital. *Revista científico profesional de la pedagogía y psicopedagogía, 5,* 113-127.

Choi, S., & Mao, X. (2021). Teacher autonomy for improving teacher self-efficacy in multicultural classrooms: A cross-national study of professional development in multicultural education. *International Journal of Educational Research*, *105*, 101711. https://doi.org/10.1016/j.ijer.2020.101711

Colson, T., Xiang, Y., & Smothers, M. (2021). How professional development in co-teaching impacts self-efficacy among rural high school teachers. *Rural Educator*, *42*(1), 20-31. https://doi.org/10.35608/ruraled.v42i1.897

Corujo-Vélez, C., Barragán-Sánchez, R., Hervás-Gómez, C., & Palacios-Rodríguez, A. (2021). Teaching innovation in the development of professional practices: Use of the collaborative blog. *Education Sciences, 11*(8), 1-18. http://doi.org/10.3390/educsci11080390

Cypress, B.S. (2017). Rigor or Reliability and Validity in Qualitative Research: Perspectives, Strategies, Reconceptualization, and Recommendations. *Dimensions of Critical Care Nursing 36*(4), 253-263, https://doi.org/10.1097/DCC.0000000000000253

Recibido: 26/03/2023 Aceptado: 03/05/2024

Eurydice (2017). *The Teaching Profession in Europe: Practices, Perceptions, and Policies. Eurydice Report*. Publications Office of the European Union.

Fernández-Sánchez, H., King, K., & Enríquez-Sánchez, C.B. (2020). Revisiones Sistemáticas Exploratorias como metodología para la síntesis del conocimiento científico. *Enferm. Univ, 17*(1), 87-94. https://doi.org/10.22201/eneo.23958421e.2020.1.697

Flick, U. (2018). *An introduction to qualitative research*. SAGE.

Gast, I., Schildkamp, K., & van der Veen, J.T. (2017). Team-Based Professional Development Interventions in Higher Education: A Systematic Review. *Review of Educational Research, 87*(4), 736-767. https://doi.org/10.3102/0034654317704306

Grau, S., Gómez, C., & Perandones, T. (2009). La formación del profesorado como factor decisivo de la excelencia educativa. En C. Gómez y S. Grau (Coords.) *Propuestas de diseño, desarrollo e innovaciones curriculares y metodología en el EEES* (pp. 7-26). Universidad de Alicante.

Hattie, J. (2015). *What Doesn't Work in Education: The Politics of Distraction*. Pearson.

Ibrahim, R., Norman, H., Nordin, N., & Azwar, K. (2020). Research trends in ICT and education: A systematic review of continuous professional development courses for online educators. *Journal of Physics: Conference Series*, *1529*, 1-7. http://doi.org/10.1088/1742-6596/1529/4/042060

Irby, B., Pashmforoosh, R., Druery, D.M., Eljaouhari, N., Tong, F., Lara-Alecio, R. (2022). An Analysis of Virtual Professional Development for School Leaders During COVID-19. *International Journal of Virtual and Personal Learning Environments, 12*(1), 1-19. https://doi.org/10.4018/IJVPLE.302097

Lindvall, J., & Ryve, A. (2019). Coherence and the positioning of teachers in professional development programs. A systematic review. *Educational Research Review*, 27, 140-154. https://doi.org/10.1016/j.edurev.2019.03.005

Lo, N. P. K., & To, B. K. H (2023). The transformation of identity of secondary school teachers: Professional development and English language education strategies in Hong Kong during the COVID-19 pandemic. *Cogent Education, 10*(1), 1-32. https://doi.org/10.1080/2331186X.2022.2163790

Luo, T., Freeman, C., & Stefaniak, J. (2020). "Like, comment, and share"—professional development through social media in higher education:

A systematic review. *Education Tech Research Dev, 68*, 1659-1683. https://doi.org/10.1007/s11423-020-09790-5

Matschiner, A. (2023). A Systematic Review of the Literature on Inservice Professional Development Explicitly Addressing Race and Racism. *Review of Educational Research August, 93*(4), 594-630. https://doi.org/10.3102/00346543221125245

Ministerio de Educación, Cultura y Deporte (2015). *El profesorado del siglo XXI. XXI Encuentro de Consejos Escolares Autonómicos y del Estado*. MECD.

Mitchell, R. (2013). What is professional development, how does it occur in individuals, and how may it be used by educational leaders and managers for the purpose of school improvement? *Professional Development in Education, 39*(3), 387-400. http://dx.doi.org/10.1080/19415257.2012.762721

Morgan, H. (2022). Understanding Thematic Analysis and the Debates Involving Its Use. *The qualitative report*, *27*(10), 2079-2091. https://doi.org/10.46743/2160-3715/2022.5912

Morin, K. L., Ganz, J. B., Vannest, K. J., Haas, A. N., Nagro, S. A., Peltier, C. J., & Ura, S. K. (2019). A systematic review of single-case research on video analysis as professional development for special educators. *The Journal of Special Education*, *53*, 3-14.

Newman, M., & Gough, D. (2020). Systematic Reviews in Educational Research: Methodology, Perspectives and Application. En O. Zawacki-Richter et al. (Eds.) *Systematic Reviews in Educational Research. Methodology, Perspectives and Application* (pp. 3-22). Springer.

Nieva, J. A. y Martínez, O. (2016). Una nueva mirada sobre la formación docente. *Universidad y Sociedad, 8*(4), 14-21.

Nunn, J., & Chang, S. (2020). What are Systematic Reviews? *WikiJournal of Medicine, 7*(5), 1-1. https://doi.org/10.15347/wjm/2020.005

OCDE (2021). *Teachers Getting the Best out of Their Students: From Primary to Upper Secondary Education*, TALIS. OECD Publishing. https://doi.org/10.1787/5bc5cd4e-en

OCDE (2023). *Review education policies. Teacher professional development*. OECD Publishing. https://gpseducation.oecd.org/revieweducationpolicies/#!node=41732&filter=all

Pamies, Ma, Cascales, A., & Gomariz, M.A. (2022). Association between conditioning factors of the transfer oflifelong learning and the

application of learning in the training programmes of Pre-school and Primary Education. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, *25*, 127-40.

Page, M. et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*(71), 1-9. http://dx.doi.org/10.1136/bmj.n71

Pedroza, R., & Reyes, A.M. (2019). Formación docente autorreflexiva para el aprendizaje en matemáticas. *V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad (CINAIC 2019),* Universidad Politécnica de Madrid, Madrid.

Perry, E. (2023). Teacher Professional Development in Changing Circumstances: The Impact of COVID-19 on Schools' Approaches to Professional Development. *Education Sciences, 13*(48). https://doi.org/10.3390/educsci13010048

Poulou, M. S., Reddy, L. A., & Dudek, C. M. (2023). Teachers and school administrators' experiences with professional development feedback: The classroom strategies assessment system implementation. *Frontiers in Psychology*, *14*, 1074278. https://doi.org/10.3389/fpsyg.2023.1074278

Pozdeeva, S.I. (2018). Participation of primary school teachers in educational innovations as the groundwork for their professional development: Organization and management. *Advances in Intelligent Systems and Computing, 677*, 28-36. http://doi.org/10.1007/978-3-319-67843-6_4

Reychav, I., Elyakim, N., & McHaney, R. (2023). Lifelong learning processes in professional development for online teachers during the Covid era. *Front. Educ*, *8*, 1041800. https://doi.org/10.3389/feduc.2023.1041800

Rojas, N., & Martínez, O. (2023). Formación del Docente y Pedagogía de la Complejidad en el Contexto de la Educación. *Revista Scientific, 1*(1), 249-265. https://doi.org/10.29394/scientific.issn.2542-2987.2016.1.1.14.249-265

Santos, M.A. (2020). La formación del profesorado. Piedra angular del sistema educativo. *Crónica, Revista científico profesional de la pedagogía y psicopedagogía,* 5, 81-87.

Savard, A., & Cyr, S. (2018). A waterfall model for providing professional development for elementary school teachers: A pilot project to

implement a competency-based approach. *Global Education Review, 5*(3), 165-182.

Suárez, N., Mena, D., Gómez, V., & Fernández, A.I. (2019). *La formación del profesorado en Iberoamérica. Tendencias, reflexiones y experiencias*. Universidad Tecnológica Indoamérica.

Surahman, E., & Wang, T-H. (2023). In-service STEM teachers' professional development programmes: A systematic literature review 2018–2022. *Teaching and Teacher Education*, 135, 104326. https://doi.org/10.1016/j.tate.2023.104326

Vaillant, D. (2016). El fortalecimiento del desarrollo profesional docente: una mirada desde Latinoamérica. *Journal of Supranational Policies of Education*, *5*, 5-21. http://dx.doi.org/10.15366/jospoe2016.5

Van Grieken, C. (2016). Formación docente y escuela: posibles articulaciones. *Ciencia y Sociedad, 41*(1), 29-43. https://doi.org/10.22206/cys.2016.v41i1.pp029-043

Wächter, T., & Gorges, J. (2023). How does inclusion come into schools? Attitude towards inclusion as a predictor of teachers' inclusion-related motivation for professional development courses. *Zeitschrift fur Erziehungswissenschaft, 26,* 31-54. https://doi.org/10.1007/s11618-022-01129-5

Westbroek, H., De Vries, B., Walraven, A., Handelzalts, A., & McKenney, S. (2019). Teachers as co-designers: Scientific and colloquial evidence on teacher professional development and curriculum innovation. En J. Pieters, J. Voogt y N. Pareja (Eds.) *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning* (pp.35-54). Springer.

Wu, W., Hu, R., Tan, R., & Liu, H. (2022). Exploring Factors of Middle School Teachers' Satisfaction with Online Training for Sustainable Professional Development under the Impact of COVID-19. *Sustainability*, *14*, 1-18. https://doi.org/10.3390/su142013244

Zeng, J. (2023). A theoretical review of the role of teacher professional development in EFL students' learning achievement. *Heliyon*, 9, e15806. https://doi.org/10.1016/j.heliyon.2023.e15806

Información de contacto: Universidad Complutense de Madrid, Facultad de Educación, Departamento de Investigación y Psicología. Calle Rector Hoyo Villanova s/n, 28040, Madrid. E-mail: ccamilli@ucm.es

Recibido: 26/03/2023 Aceptado: 03/05/2024

Anexos

- 68 artículos finales https://bit.ly/3TOecwv
- Base de datos (cribado documentos) https://bit.ly/3O7wIhO

Recibido: 26/03/2023 Aceptado: 03/05/2024

Validación española del Instrumento de Vinculación Estudiantil en Formación Profesional

Spanish validation of the Vocational Engagement Instrument in Vocational Training

https://doi.org/10.4438/1988-592X-RE-2024-406-644

Inmaculada Sureda-García
https://orcid.org/0000-0002-5024-1799
Universitat de les Illes Balears, Institute of Research and Innovation in Education o Institut de Recerca I Innovació

Rafael Jiménez
https://orcid.org/0000-0002-1181-3221
Universitat de les Illes Balears, Instituto de Investigaión Sanitaria Illes Balears (IdISBa)

Albert Sesé
https://orcid.org/0000-0003-3771-1749
Universitat de les Illes Balears, Instituto de Investigaión Sanitaria Illes Balears (IdISBa)

Francesca Salvà-Mut
https://orcid.org/0000-0001-8517-0014
Universitat de les Illes Balears, Institute of Research and Innovation in Education o Institut de Recerca I Innovació

Resumen

El concepto de vinculación escolar desempeña un papel crucial en la comprensión de los factores de riesgo que conducen al abandono escolar, especialmente entre un alumnado vulnerable. La literatura ha desglosado la vinculación escolar en tres subtipos: aspectos académicos, cognitivos y emocionales (Fredericks et al., 2004). Este modelo se revela de suma importancia en la contribución a estudios empíricos que evalúan la vinculación académica entre estudiantes de formación profesional, dado que este nivel educativo enfrenta un índice significativo de abandono. En este contexto, este artículo

presenta pruebas de validez de las puntuaciones del Instrumento de Vinculación Estudiantil (IVE), una nueva herramienta psicométrica diseñada para medir la Vinculación del Alumnado en el ámbito de la Formación Profesional (FP). El estudio se llevó a cabo utilizando una muestra estratificada por conglomerados compuesta por 4522 estudiantes (64,6% hombres) pertenecientes a tres regiones españolas: Mallorca (n = 1511, 33.4%), Barcelona (n = 1038, 23%), y Valencia (n = 1973, 43.6%). Los participantes estaban matriculados en el primer año de FP de grado básico (n = 1370, 30.3%) o FP de grado medio (n = 3152, 69.7%). Se aplicó una estrategia de comparación de modelos que incorporó diversas estructuras latentes, como unifactorial, pentafactorial y bifactorial, mediante Análisis Factorial Confirmatorio (CFA) y modelos de Ecuaciones Estructurales Exploratorias (ESEM), utilizando el método de estimación WLSMV. El modelo bifactorial demostró el mejor ajuste (ji-cuadrado/gl=4.64, RMSEA=.031, CFI=.970, TLI=.970, SRMR=.039). Además, se verificó la invariancia de medias considerando el nivel de FP (básico/grado medio) (RMSEA=.028 y CFI=.910) y el sexo autoinformado (masculino/femenino) (RMSEA=.029, CFI=.913). Estos resultados prometedores sobre la estructura latente del IVE pueden conducir a la identificación de patrones diferenciales de vinculación estudiantil. Asimismo, podrían facilitar el desarrollo e implementación de programas de intervención centrados en la minimización del abandono escolar y en la mejora del rendimiento académico.

Palabras clave: vinculación estudiantil, formación profesional, validación de test, redes psicométricas, modelos latentes.

Abstract

The concept of student engagement plays a crucial role in understanding the risk factors leading to dropout, particularly among vulnerable students. The literature has divided school engagement into three subtypes: academic, cognitive, and emotional (Fredericks et al., 2004). This model proves highly relevant in contributing to empirical studies that assess academic engagement among vocational education students, as this educational level faces a significant dropout rate. In this context, this article presents evidence of the validity of scores from the Vocational Engagement Instrument (VEI), a new psychometric tool designed to measure engagement in the context of Vocational Education and Training (VET) students. The study was conducted using stratified cluster sampling of 4,522 students (64.6% males) from three Spanish regions: Mallorca (n = 1,511, 33.4%), Barcelona (n = 1,038, 23%), and Valencia (n = 1,973, 43.6%). Participants were enrolled in the first year of basic VET (n = 1,370, 30.3%) or intermediate VET (n = 3,152, 69.7%). A model comparison strategy included different latent structures (unifactorial, pentafactorial, bifactor) through Confirmatory Factor Analysis (CFA) and Exploratory Structural Equation Models (ESEM), using the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimation method. The bifactor

Recibido: 31/03/2023 Aceptado: 05/12/2023

model achieved the best overall fit (chi-square/df=4.64, RMSEA=.031, CFI=.970, TLI=.970, SRMR=.039). Furthermore, measurement invariance was verified considering the VET level (basic/intermediate) (RMSEA=.028 and CFI=.910) and self-reported sex (male/female) (RMSEA=.029, CFI=.913). These promising results regarding the latent structure of the VEI may lead to the identification of differential patterns of student engagement. Additionally, they could facilitate the development and implementation of intervention programs focused on minimizing school dropout and enhancing academic performance.

Keywords: student engagement, vocational education training, test validation, psychometric networks, latent models.

Introducción

La literatura sobre el abandono escolar temprano se ha centrado en el estudio de los factores de riesgo, las interacciones entre estos factores y los procesos que conducen al abandono escolar (Soler et al., 2021). En línea con esta perspectiva, Reschly y Christenson (2012), a través de su programa de intervención "Check and Connect" (diseñado para promover la participación de los estudiantes a través de la construcción de relaciones, la resolución de problemas y la persistencia de los estudiantes marginados) establecen un modelo que tiene como objetivo relacionar el contexto, la participación del alumnado y sus resultados académicos. Este modelo es de especial relevancia para los objetivos del presente estudio con estudiantes de FP y representa una contribución relevante debido a la escasez de información y estudios empíricos sobre este tema y etapa formativa en la educación española (Echeverría y Martínez, 2017). En el sistema educativo español, la educación secundaria profesional se organiza en dos niveles: FP de grado básico (en adelante FPGB) y FP de grado medio (en adelante FPGM). La FPGB está dirigida a estudiantes que no han finalizado la educación secundaria obligatoria. La edad de ingreso es de 15 años, y tiene una duración de dos cursos académicos. La literatura sobre las especificidades de este alumnado explica la vulnerabilidad social y educativa como característica personal de este colectivo, mostrando los altos niveles de abandono en esta formación (Olmos et al., 2020). En la FPGM, los estudiantes tienen el título de *GESO* o formación equivalente, sin embargo, el acceso a estos estudios está infravalorado en comparación con los estudios de

bachillerato español. Estos estudiantes provienen de diversas formas de admisión, incluyendo itinerarios educativos no lineales (aquellos que han experimentado interrupciones en sus estudios académicos de secundaria o postsecundaria, o en el mercado laboral) (Cerdà et al., 2022).

A pesar de la falta de estudios específicos, dentro de este marco teórico, el concepto de vinculación estudiantil juega un papel fundamental en la comprensión de los procesos de abandono en la FP y la prevención de dicho abandono en otros estudios académicos (OCDE, 2023). El constructo vinculación estudiantil se refiere a que el alumnado esté involucrado o comprometido con las actividades académicas y sociales de su escuela (Wang et al., 2019). El concepto de vinculación ha surgido como una forma de comprender y mejorar los resultados de los y las estudiantes en riesgo de fracaso escolar.

Factores del Instrumento de Vinculación Estudiantil (IVE)

La literatura concluyó que la vinculación del alumnado se componía de tres subtipos: conductual-académico (p. ej., conducta positiva, esfuerzo, participación), cognitivo (p. ej., autorregulación, metas de aprendizaje, inversión en el aprendizaje) y emocional o afectivo (p. ej., pertenencia y relaciones con el profesorado, compañeros y compañeras de clase y familia) (Fredericks et al., 2004). En base al trabajo teórico de Reschly y Christenson (2012), Fredricks et al. (2004) y la teoría cognitivo social del desarrollo de la carrera que refuerza el estudio de variables cognitivas como la autoeficacia y las expectativas de resultados (Lent et al.,1994), se proponen cinco factores en el IVE. Los factores son los siguientes: deberes escolares y disciplina, metas y expectativas, integración social (iguales del grupo clase), apoyo del profesorado y apoyo de la familia.

En primer lugar, los deberes y la disciplina escolar (SD) explican la vinculación conductual en toda la literatura (Estell & Perdue, 2013; Fredericks et al., 2004)). Los deberes escolares y la disciplina reflejan como el alumnado se enfoca en hacer la tarea académica para la evaluación de comportamientos como la persistencia, el esfuerzo y la atención del estudiante. Además, este factor contempla conductas positivas, como seguir las reglas, adherirse a las normas del aula y la ausencia de comportamientos disruptivos como faltar a la escuela o meterse en problemas.

Recibido: 31/03/2023 Aceptado: 05/12/2023

Un segundo factor identifica las metas y expectativas (GE) del alumnado, que implican la vinculación cognitiva. El concepto de expectativas de resultados implica las consecuencias anticipadas de un curso de acción. La expectativa de resultados vocacionales podría considerarse como las consecuencias imaginarias de la realización de conductas académicas y profesionales que serían útiles para las opciones y decisiones profesionales posteriores (Betz y Voyten, 1997).

El tercer factor explica la integración social (SI) con los compañeros y compañeras de clase. El tipo de relación y el apoyo social recibido por el grupo de pares condiciona la participación estudiantil a corto y largo plazo del alumnado. Una buena relación con los pares proporciona apoyo emocional, mejora la autoestima y el desarrollo de la identidad, y tiene importantes consecuencias para el bienestar social, emocional y cognitivo de los sujetos (Sureda et al., 2021). El cuarto factor determina el apoyo del profesorado (TS) como una clave importante en la vinculación emocional, ya que se considera que el profesorado tiene una influencia próxima y, por tanto, crucial en la vinculación del alumnado (Niittylahti et al., 2019; Quin et al., 2018). Finalmente, el quinto factor contempla el apoyo de la familia (FS). El alumnado que cuenta con el apoyo y el compromiso percibido de sus familias muestra más interés y se involucra más en las actividades académicas, evitando así la desvinculación escolar o el riesgo de abandono escolar (Sureda et al., 2021; Wang et al., 2019).

Instrumentos Relevantes para el IVE

El nuevo instrumento abordó la concepción multicomponente de la vinculación, dirigida a una intervención práctica para mejorar la vinculación del alumnado, y se dirigió a un perfil de estudiantes en formación profesional. En este caso, los tres instrumentos más relevantes por los criterios mencionados y por el diseño y construcción de nuestra medida de evaluación son el *Student Engagement Instrument* (SEI) (Appleton et al., 2006), la escala *Trousse d'évaluation des décroheurs potentiels* (TEDP) (Janosz et al., 2007) y la *Vocational Outcome Expectations Scale* (VOE; McWhirter et al., 2000).

El *Student Engagement Instrument* (SEI), utilizado en el proyecto "Check & Connect" (Appleton y Christenson, 2004; Appleton et al., 2006), operacionaliza la vinculación emocional y cognitiva, utilizando 35 ítems.

Recibido: 31/03/2023 Aceptado: 05/12/2023

Aborda diferentes contextos de interacción (escuela, familia y comunidad) y los resultados del alumnado (vinculación, éxito y abandono para evaluar la vinculación conductual) de manera conjunta y relacionada. Este instrumento se basa en variables altamente modificables a partir de la intervención escolar, posibilitando acciones educativas dirigidas a prevenir y reducir el abandono escolar (Reschly y Christenson, 2012).

Este instrumento utilizó una muestra inicial de 1931 estudiantes de noveno grado de un distrito escolar urbano étnicamente diverso, mayoritariamente de bajos ingresos. Otros estudios sobre SEI utilizaron una muestra de estudiantes de 6º a 12º grado. En cuanto a las propiedades psicométricas, se obtuvieron evidencias de validez de criterio con correlaciones significativas con las variables académicas (rendimiento en lectura y matemáticas, número de asignaturas suspendidas y promedio de calificaciones). El instrumento también obtuvo evidencia adecuada de consistencia interna de sus subescalas que van de .72 a .92. Este comportamiento psicométrico adecuado también se complementó con una adecuada fiabilidad test-retest, que osciló entre .60 y .62. (Christenson et al., 2012).

Entre las limitaciones del SEI, se señala que en algunos ítems es un poco complicado establecer si sus contenidos son mutuamente excluyentes y, por lo tanto, afirmar su asignación a un solo subtipo de vinculación (Mameli y Passini, 2017). Sin embargo, el SEI debe completarse con datos académicos de la administración educativa u otros instrumentos, que nos permitan conocer la vinculación conductual con el fin de evaluar la vinculación multidimensional. Además, el SEI fue diseñado para atender a una población de educación secundaria general y no, específicamente, a una población de educación secundaria vocacional.

Otra herramienta relevante es el *Trousse d'évaluation des décroheurs potentiels* (TEDP) (Janosz et al., 2007), dirigido a la práctica educativa, y diseñado para predecir el abandono escolar en diferentes estudios académicos. El TEDP se compone de 54 ítems, utilizados originalmente en Quebec, y permite identificar tipologías de estudiantes, con edades comprendidas entre los 12 y los 18 años, que corren el riesgo de abandonar la escuela antes de graduarse en la educación secundaria. Las respuestas se ordenan de acuerdo con 4 opciones de tipo Likert: 1, falso; 2, algo falso; 3, algo cierto; 4, verdadero; aunque el contenido de algunas preguntas puede modificar el formato de la respuesta y ofrecer otras opciones, o una respuesta abierta.

Recibido: 31/03/2023 Aceptado: 05/12/2023

El estudio de validación del instrumento TEDP proporciona evidencias adecuadas de validez de decisión, con análisis de la curva ROC, y valores de área bajo la curva apropiados entre .70 y .85. En cuanto a la estabilidad en el tiempo de la medida, se obtienen valores adecuados en torno a .70. La validación se llevó a cabo en una muestra de 35000 estudiantes de secundaria, con edades comprendidas entre los 12 y los 18 años, en 79 centros educativos evaluados como parte de una estrategia de intervención (Janosz et al., 2007). El TEDP indica que, aunque los factores de riesgo de abandono escolar son numerosos y tienen múltiples orígenes (factores individuales, familiares, socioeconómicos, etc.), no es necesario medirlos todos para obtener una evaluación fiable (Janosz et al., 2007). Por otro lado, permite postular que no todos los que abandonan de manera temprana la escuela tienen las mismas vulnerabilidades o fortalezas, ya que su experiencia escolar o social puede ser diferente. Entre sus limitaciones, mide la vinculación escolar de acuerdo con el instrumento *l'Adaptation Sociale et Personnelle pour les Adolescents Québécois* (MASPAQ) de LeBlanc (1996), centrándose más en los factores conductuales de la vinculación. Entre otras limitaciones de este instrumento, el índice de riesgo de abandono varía en función de los estudios que se cursan y de las características de los estudiantes (edad, sexo, formación), por lo que no existe, a priori, un índice de riesgo de abandono.

El tercer instrumento de interés para el estudio del perfil del alumnado de formación profesional es el *Vocational Outcome Expectations Scale* (VOE) (McWhirter et al., 2000). Específicamente, las expectativas de resultados vocacionales podrían considerarse como las consecuencias imaginarias de realizar conductas académicas y profesionales que serían útiles para las opciones y decisiones profesionales posteriores en sus estudios académicos secundarios o postsecundarios. (Betz y Voyten, 1997; Işik, 2013). El VOE es una escala de 12 ítems que mide el nivel de expectativas positivas de las personas encuestadas con respecto a los resultados de su elección de carrera. Las calificaciones se realizan en una escala de 4 puntos con anclajes que van desde 1 "totalmente en desacuerdo" hasta 4 "totalmente de acuerdo". El rango de puntuaciones posibles varía desde una puntuación mínima de 12 hasta una puntuación máxima de 48, y las puntuaciones más altas reflejan expectativas de resultados más positivas. La validez concurrente de la escala fue apoyada por una correlación positiva (r = .54) con otra escala de expectativas de

Recibido: 31/03/2023 Aceptado: 05/12/2023

resultados (Fouad y Smith, 1996). McWhirter et al. (2000) demostraron que la fiabilidad test-retest de la escala era de .59, y que el coeficiente de consistencia interna Alfa de Cronbach era de .83. Işik (2013) determinó que la fiabilidad test-retest de la versión turca fue de .79 y el coeficiente de consistencia interna Alfa de Cronbach fue de .87, lo que indica una fiabilidad adecuada. Las limitaciones de esta escala, en términos de nuestros objetivos, son que no utiliza una medida multicomponente de la vinculación de los estudiantes. Sobre la base de las fortalezas de los instrumentos revisados, el objetivo fue validar las puntuaciones de un nuevo instrumento que permita evaluar la vinculación del alumnado en una población de FP.

Métodos

Participantes

La muestra estuvo formada por 4522 estudiantes españoles (64.6% varones), procedentes de Mallorca (n = 1511, 33.4%), Barcelona (n = 1038, 23%) y Valencia (n = 1973, 43.6%), todos ellos matriculados en el primer curso de FPGB (n = 1370, 30.3%) o FPGM (n = 3152, 69.7%) (Tabla I). Se utilizó un muestreo estratificado por conglomerados. Los estratos se formaron a partir de 17 familias profesionales según el tipo de centro y la ubicación. Esta población inicial se restringió a estudiantes de 14 a 19 años con el fin de asegurar la homogeneidad evolutiva de la muestra. En el caso del alumnado de FPGB, representaron una edad media de 16.06 años (*SD* = 0.77) y en el caso de la FPGM de 17.34 años (*SD* = 0.97). El conjunto de la muestra representó una edad media de 16.88 años (*SD* = 1.09).

Desarrollo del instrumento

El nuevo instrumento considera la importancia de evaluar factores sobre la dimensión emocional, cognitiva y conductual de la vinculación estudiantil (como la integración social con los compañeros, el apoyo del profesorado, el apoyo de la familia, la meta profesional, y los deberes

Recibido: 31/03/2023 Aceptado: 05/12/2023

TABLA I. Descripción de la muestra

		FPGB		FPGM		Total	
		Recuento	**%**	**Recuento**	**%**	**Recuento**	**%**
Región	Mallorca	354	25.8	1157	36.7	1511	33.4
	Barcelona	277	20.2	761	24.1	1038	23.0
	Valencia	739	53.9	1234	39.1	1973	43.6
Sexo	Masculino	975	71.6	1937	61.6	2912	64.6
	Femenino	387	28.4	1206	38.4	1593	35.4
Inmigrante	Sí	167	26.7	456	24.0	623	24.6
	No	459	73.3	1447	76.0	1906	75.4
Intención de abandonar la escuela	Sí	285	21.3	756	24.4	1041	23.4
	No	1053	78.7	2348	75.6	3401	76.6
Expulsiones de la escuela	Sí	630	46.8	672	21.6	1302	29.2
	No	717	53.2	2438	78.4	3155	70.8
Aprobar el curso sin asistir a clases	Sí	235	17.9	341	11.3	576	13.3
	No	1080	82.1	2684	88.7	3764	86.7
Número de cursos repetidos	0	130	9.8	1231	40.2	1361	31.0
	1	753	56.8	1254	40.9	2007	45.7
	2	386	29.1	453	14.8	839	19.1
	3	35	2.6	70	2.3	105	2.4
	4	20	1.5	44	1.4	64	1.5
	> 4	1	0.1	12	0.4	13	0.3

Fuente: elaboración propia.

y disciplina escolar) por parte del SEI (Appleton y Christenson, 2004) y el TEDP (Janosz et al., 2007), incorporando la evaluación de las expectativas creadas hacia los estudios como una variable significativa en los estudios vocacionales desde el instrumento VOE (McWhirter et al., 2000) (Tabla II).

El IVE tenía inicialmente 52 ítems con cuatro opciones de respuesta tipo Likert (totalmente en desacuerdo, en desacuerdo, de acuerdo y totalmente de acuerdo) que operacionalizan una estructura latente de cinco dimensiones. El contenido de los ítems y sus estadísticas descriptivas se presentan en una tabla que figura en el Anexo.

TABLA II. Factores utilizados en el IVE

Instrumentos relevantes en el IVE	**Factores utilizados en el IVE**
Student Engagement Instrument (SEI)	Control y pertinencia del trabajo escolar y de las aspiraciones y metas futuras (10 ítems)
	Relaciones profesor-alumno (9 ítems)
	Apoyo entre iguales al aprendizaje (5 ítems)
	Apoyo familiar (13 ítems)
Trousse d'évaluation des décroheurs potentiels (TEDP)	Deberes escolares y disciplina escolar (10 ítems)
	Percepción del compromiso de los padres (9 ítems)
Vocational Outcome Expectations Scale (VOE)	Expectativas profesionales (5 ítems)

Fuente: elaboración propia.

Validez del contenido

El instrumento fue evaluado por un panel de 16 personas expertas (9 mujeres) con experiencia en estudios profesionales, secundarios y universitarios para la validación de contenidos (seis profesores de formación profesional, cinco de educación secundaria y cinco de universidades españolas). De acuerdo con Poli y Hungler (2000), se establecieron dos niveles de análisis de los ítems en función de su grado de relevancia y pertinencia utilizando valores de 1 (no relevante o no pertinente) a 4 (muy relevante o pertinente). También se solicitaron sugerencias sobre cuestiones no pertinentes o no relevantes en este contexto o población. Las personas expertas podían hacer una propuesta de cambio en la redacción o la eliminación de los ítems. Finalmente, haciendo un trabajo grupal presencial con los expertos, se concretaron las sugerencias, valoraciones y acuerdos. Las medias estadísticas de concordancia fueron de 3.65 para la relevancia y 3.68 para la pertinencia. El valor medio más bajo fue de 3 con *V* de Aiken = .67 e IC del 90% de la *V* de Aiken = [.55;.77], y 4 el valor más alto con *V* de Aiken = 1 e IC del 90% de la *V* de Aiken = [.95;1]. La relevancia y pertinencia de los ítems con el IC de la *V* de Aiken, así como los cambios en la redacción o eliminación de los ítems, se proporcionan en el siguiente archivo

Recibido: 31/03/2023 Aceptado: 05/12/2023

recuperado del repositorio académico en línea de Zenodo: https://doi.org/10.5281/zenodo.10257332.

Validez aparente

El cuestionario se aplicó a una muestra piloto de 172 estudiantes de FP (24.4% en FPGB y 75.6% en FPGM, 65.1% varones, *M* = 18.4 años). Se les pidió que completaran el cuestionario por escrito. Posteriormente, se les animó a hacer comentarios sobre la formulación de las preguntas y si el cuestionario atendía a su comprensión de la vinculación estudiantil. Basándonos en el alfa de Cronbach ordinal (Elosua y Zumbo, 2008), algunos ítems del instrumento original fueron reformulados a partir de los comentarios proporcionados por el alumnado participante en el estudio piloto, mejorando la traducción de los ítems sin pérdida de validez de contenido. En general, los ítems fueron entendidos como intencionados y fue posible responder con las alternativas de respuesta ofrecidas. En el estudio piloto se calcularon los valores del coeficiente alfa ordinal, y todos los factores obtuvieron una consistencia interna aceptable con valores que oscilaron entre .76 y .81.

Procedimiento

El estudio fue aprobado por el Comité de Ética de la Investigación de la Universitat de les Illes Balears (nº de registro 5488), lo que permitió crear un fichero cifrado de datos personales gestionado por la universidad y el grupo de investigación. Tras obtener la colaboración de los centros, se realizaron reuniones informativas con la dirección de los centros y, posteriormente, se obtuvo el consentimiento informado de las familias del alumnado menor de edad. Se informó a los estudiantes que su participación sería anónima, confidencial y voluntaria. El tiempo medio necesario para cumplimentar el cuestionario fue de 30 minutos. La encuesta se realizó durante un período de dos a tres meses después del inicio del año escolar, antes de los primeros resultados académicos, evitando que la evaluación académica interfiriera con las respuestas, así como la pérdida (más adelante) de una parte del alumnado. El instrumento se administró dentro de una clase regular sin introducir factores de distracción.

Recibido: 31/03/2023 Aceptado: 05/12/2023

Análisis de datos

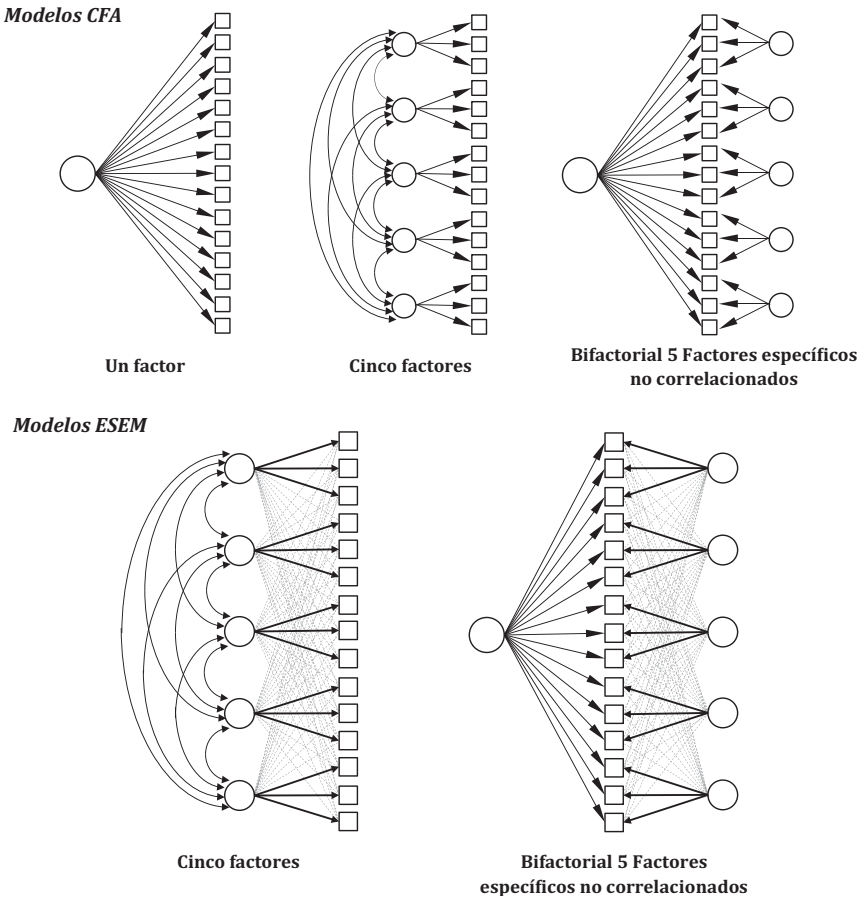
Se implementó el análisis descriptivo univariante de los datos para explorar el comportamiento estadístico de los 52 ítems, evaluar la calidad de los datos y verificar el cumplimiento de los supuestos estadísticos (normalidad univariante y multivariante). No se implementaron métodos de imputación para los valores faltantes, utilizando la eliminación por pares. La ocurrencia de valores faltantes en las respuestas de los ítems fue prácticamente insignificante, con su presencia porcentual fluctuando entre el 0% y 1,2%. De manera preliminar, los 52 ítems de la versión inicial del cuestionario fueron sometidos a un análisis de estructura latente mediante Redes Psicométricas (PN, por sus siglas en inglés). Se utilizó un Análisis Gráfico Exploratorio (EGA, por sus siglas en inglés) (Golino & Epskamp, 2017) mediante el paquete EGAnet (Hudson, 2020) del programa R. En concreto, se utilizaron los métodos "glasso" y "TMFG" (*Triangulated Maximally Filtered Graph*) para estimar la red latente (Massara et al., 2017). Estos métodos permiten detectar la existencia de factores latentes a partir de agrupaciones de ítems en una estructura de red interrelacionada. Se implementaron los algoritmos de estimación "walktrap" y "louvain". Se realizó una simulación mediante bootstrap paramétrico (1000 muestras) para evaluar la estabilidad de los modelos obtenidos. Con esta estrategia analítica se pretendió un doble objetivo: comprobar si las PN captan la hipotética estructura teórica de cinco factores, y detectar y eliminar ítems con mal funcionamiento psicométrico en la estructura latente.

A continuación, se utilizó esta versión refinada que implementa las PN para llevar a cabo una estrategia de comparación de modelos que incluía 5 modelos hipotéticos que compiten entre sí: modelos de un factor, cinco factores y bifactorial bajo estimación de CFA (Análisis Factorial Confirmatorio); y para descartar la existencia de altas cargas cruzadas de ítems, se aplicó el Modelo de Ecuaciones Estructurales Exploratorias (ESEM, por sus siglas en inglés) tanto al modelo de cinco factores como al bifactorial. El diagrama conceptual de todos los modelos se presenta en la Figura I.

Todos los modelos propuestos se estimaron mediante el método de Mínimos Cuadrados Ponderados con Media y Varianza ajustados (WLSMV, por sus siglas en inglés) con errores estándar robustos, ya que la escala Likert de los ítems tenía menos de cinco categorías de

Recibido: 31/03/2023 Aceptado: 05/12/2023

FIGURA I. Diagrama conceptual de los modelos de estructura latente que compiten entre sí



Fuente: elaboración propia.

respuesta. El método WLSMV emplea la matriz de correlación policórica al realizar un análisis factorial confirmatorio con variables categóricas. Este enfoque es particularmente adecuado para situaciones en las que trabajar con datos que se desvían del supuesto de normalidad involucra variables de naturaleza ordinal o nominal (Brauer et al., 2023; Martínez-Abad y Rodríguez-Conde, 2017; Muthén et al., 1997). Una vez obtenido el modelo mejor ajustado, se implementaron análisis adicionales de

invarianza multigrupo en dos variables: nivel de estudios de FP (Básico/Grado medio) y sexo autoinformado (Hombre/Mujer). Se utilizó un procedimiento de medición de la invarianza (Brown, 2014) que incluía cinco configuraciones de invarianza, desde el nivel más bajo hasta el más alto de restricciones de igualdad en los parámetros del modelo a estimar entre diferentes submuestras. El primer paso fue la invarianza configuracional, que examina si los datos de las dos submuestras están representados por la misma estructura factorial. Se estimó el segundo paso, la invarianza débil/métrica, en la que las cargas factoriales se restringen a ser equivalentes entre subgrupos. Se completó el tercer paso, la invariancia fuerte/escalar, en la que todas las intercepciones se restringen a ser iguales en las dos submuestras. En el cuarto paso se comprobó la invarianza estricta, en la que la varianza del error de medición se restringe a ser igual entre los grupos comparados. Y, por último, también se estimó el quinto paso, en el que se restringe la igualdad de las medias de los factores latentes.

Se utilizaron los siguientes índices para evaluar la bondad de ajuste general de los modelos: (a) el estadístico ji-cuadrado (χ^2), (b) el cociente ji-cuadrado a grados de libertad (χ^2/df), (c) RMSEA y su intervalo de confianza (IC) del 90%, con un valor p para RMSEA < .05, (d) CFI, e) TLI, y f) SRMR. Se utilizaron los valores de corte propuestos por Byrne (2016), Hair et al. (2008), Hu y Bentler (1999), Kline (2005) y Steiger (2007) para determinar un ajuste adecuado: valores de ji-cuadrado no significativos; χ^2/df ≤ 5; RMSEA ≤ .05; CFI y TLI ≥ .95; y SRMR ≤ .08.

En cuanto al análisis adecuado de los modelos bifactoriales, centrados en la fuerza del factor general y la potencial unidimensionalidad, se estimaron diferentes índices específicos. El Omega Jerárquico (ω_H) (al cuadrado) representa la variancia total explicada que se puede estimar para los factores generales y específicos; cuando el valor para el factor general es ≥ .70, puede ser parcialmente indicativo de unidimensionalidad, y cuando el valor para los factores específicos es ≥ .30, estos pueden considerarse significativos. El índice H permite evaluar cuán bien una variable latente está representada por un conjunto de ítems, con un valor mínimo de adecuación igual a .70, y se puede calcular tanto para el factor general como para los específicos.

El índice PUC estima la proporción de correlaciones no contaminadas por multidimensionalidad, el ECV-SS es la Variancia Común Explicada de un factor específico con respecto a sí mismo, y el ECV-SG es el ECV de un factor específico con respecto al factor general. Se recomienda

Recibido: 31/03/2023 Aceptado: 05/12/2023

interpretar conjuntamente los valores de PUC y ECV, teniendo en cuenta que la unidimensionalidad de un modelo se puede establecer cuando se obtiene $\omega_H \geq .70$, PUC > .70 y ECV > .60 (Constantinou y Fonagy, 2019; Reise y cols., 2013). En cuanto al comportamiento de los ítems, el índice ECV-i indica qué porcentaje de la variancia real de cada ítem es explicado por el factor general, esperando valores ≥ .80 para concluir sobre una influencia significativa del factor general (Stucky y Edelen, 2015); valores más bajos de ECV-i podrían indicar cargas más altas en factores específicos. Un análisis adecuado de un modelo bifactorial requiere que todos estos índices sean interpretados conjuntamente para verificar si el factor general es suficiente para explicar la variancia de los ítems, y por lo tanto, los factores específicos son irrelevantes, o por el contrario, los factores específicos pueden considerarse significativos para la medición del constructo.

Por lo que respecta al análisis de invariancia, se realizaron pruebas de $\Delta\chi^2$ y ΔCFI para comparar el ajuste de las restricciones menos restringidas a las totalmente restringidas (igualdad de parámetros entre los grupos). Dado que la prueba de ji-cuadrado es realmente restrictiva para no ser significativa (sesgada por el tamaño de la muestra) como indicativa de un buen ajuste, un ΔCFI entre modelos menor o igual a .01 también se consideró estadísticamente no significativo (Cheung y Rensvold, 2002). Los modelos CFA se estimaron utilizando *lavaan* (Rosseel, 2012) y *semTools* (Jorgensen et al., 2021) del programa R (R Core Team, 2021), y los modelos ESEM con el programa MPlus 7.4 (Muthén & Muthén, 2017).

Resultados

Los modelos de Redes Psicométricas aplicados a los 52 ítems del cuestionario IVE mostraron una clara estructura latente de cinco factores utilizando el método TMFG (100% de las repeticiones), y el método glasso encontró un modelo de seis factores en el 81.4% de las repeticiones y un modelo de cinco factores en el 18.6% restante. En ambas soluciones, los cinco factores hipotetizados para operacionalizar el constructo pueden identificarse claramente, pero existen discrepancias en cuanto a la detección de un sexto factor. Considerando su composición (5 ítems), el sexto factor es una escisión del factor Apoyo de la familia. El contenido de estos ítems, por ejemplo, "Mis padres esperan que continúe mis estudios durante el mayor tiempo posible", o "Daría un disgusto a mis padres

si dejara el instituto/colegio", trata de medir la reacción de los padres ante ciertos comportamientos o sus expectativas, de acuerdo con la percepción del alumnado (hijos). Esta forma indirecta de medir el apoyo de la familia no parece obtener un buen comportamiento psicométrico.

Además, dos ítems del factor Metas y expectativas: "Cuando hago una actividad escolar intento comprender bien lo que estoy haciendo" y "Me comparo con mis compañeros/as para saber si voy aprendiendo al ritmo adecuado", así como un ítem del factor Deberes y disciplina escolar: "Me gusta ir al instituto/colegio", presentaron menor estabilidad en sus cargas factoriales dentro de la estructura latente mediante remuestreo bootstrap, especialmente con el método glasso. Por esta razón, estos 8 ítems fueron eliminados de la versión inicial del cuestionario. La versión refinada del cuestionario con 44 ítems se muestra en el Anexo de este documento.

El siguiente paso consistió en llevar a cabo la estrategia de comparación de modelos SEM para evaluar la estructura latente que mejor se ajustaba al cuestionario. Los cinco modelos latentes hipotetizados (Figura I), CFA y ESEM, se estimaron utilizando el método WLSMV. La Tabla 3 muestra los índices generales de bondad de ajuste para todos los modelos estimados. De acuerdo con los resultados, el peor ajuste correspondió al modelo CFA de un factor (χ^2=17637.83, df=902, p<.001, χ^2/df=19.55, RMSEA=.071, IC del 90% para RMSEA=[.070;. 071], CFI=.81, TLI=.80 y SRMR=.095), en el que ninguno de estos índices alcanza los valores de corte para un ajuste adecuado. Y el mejor fue el modelo CFA Bifactorial 5F (χ^2=3983.24, df=858, p<.001, χ^2/df=4.64, RMSEA=.031, IC del 90% para RMSEA=[.030;.032], CFI=.97, TLI=.97 y SRMR=.039), en el que χ^2/df es inferior a 5, RMSEA está claramente por debajo de .05, CFI y TLI son superiores a .95, y SRMR es inferior a .08; aunque el ji-cuadrado es estadísticamente significativo, todos los índices indican un buen ajuste.

La inclusión de los modelos ESEM en la estrategia analítica responde a la necesidad de descartar la existencia de altas cargas cruzadas factoriales entre factores. Los modelos ESEM pueden considerarse un híbrido entre el análisis factorial exploratorio (EFA) y confirmatorio (CFA), ya que permite establecer relaciones a priori como los CFA, pero al mismo tiempo, no restringe las cargas cruzadas factoriales a cero como los EFA. Los resultados obtenidos para los dos modelos ESEM estimados, cinco factores y bifactorial 5F, mostraron peores valores de ajuste que los obtenidos por los modelos CFA (excepto el modelo CFA de un factor) y

Recibido: 31/03/2023 Aceptado: 05/12/2023

TABLA III. Índices generales de bondad de ajuste para los modelos CFA y ESEM y comparación de pruebas (método WLSMV)

Modelo	χ^2	**gl**	**Sig.**	χ^2**/df**	**RMSEA (IC del 90%)**	**p(RMSEA<.05)**	**CFI**	**TLI**	**SRMR**	$\Delta\chi^2$
1F CFA	17637.83	902	<.001	19.55	.071 (.070-.071)	<.001	.81	.80	.095	
5F ESEM	12199.56	736	<.001	16.58	.059 (.058-.060)	<.001	.91	.88	.080	5438.27***
Bifactorial 5F ESEM	8454.38	697	<.001	12.13	.050 (.049-.051)	.70	.94	.92	.067	3745.18***
5F CFA	4756.71	892	<.001	5.33	.033 (.032-.034)	1.00	.97	.96	.042	3697.67***
Bifactorial 5F CFA	3983.24	858	<.001	4.64	.031 (.030-.032)	1.00	.97	.97	.039	773.47***

Fuente: elaboración propia *=p<.05 **=p<.01 ***=p<.001.

se observaron diferencias estadísticamente significativas utilizando las pruebas j-cuadrado para la comparación de modelos. Tras analizar las matrices factoriales de los modelos ESEM estimados, no se detectaron cargas cruzadas factoriales con valores altos, ya que todos ellos fueron menores a .15.

Una vez obtenido un ajuste adecuado para el modelo CFA Bifactorial 5F, se realizó un estudio de invariancia considerando dos variables: el nivel de FP (Básico/Grado medio) y el sexo autoinformado (Masculino/Femenino). En la Tabla IV se muestran los índices generales de bondad de ajuste y las pruebas de comparación para todos los niveles de invariancia restringida y ambas variables.

El estudio completo de invariancia obtuvo buenos resultados de ajuste en todos los niveles de restricción (Configuracional, Débil, Fuerte, Estricto y Medias) para el nivel de FP y el sexo autoinformado. Aunque las pruebas de $\Delta\chi^2$ solo fueron no significativas entre los niveles Débil y Configuracional, el ΔCFI se mantuvo por debajo de .01 para todos

TABLA IV. Índices generales de bondad de ajuste para el modelo CFA Bifactorial 5F bajo pruebas de invariancia para el nivel de FP (Básico/Grado medio) y para el sexo autoinformado (Masculino/Femenino)

	χ^2	gl	Sig.	RMSEA (IC del 90%)	ΔRMSEA	CFI	ΔCFI	$\Delta\chi^2$
Nivel de FP								
Configuracional	4514.10	1716	<.001	.030 (.029-.031)	-	.916	-	-
Débil	5096.60	1798	<.001	.026 (.025-.027)	.004	.926	.010	53.99
Fuerte	5535.30	1836	<.001	.027 (.026-.028)	.001	.917	.009	323.64***
Estricto	5763.10	1880	<.001	.027 (.026-.028)	.000	.914	.003	114.72***
Medias	6044.40	1886	<.001	.028 (.027-.029)	.001	.910	.004	34.67***
Sexo autoinformado								
Configuracional	4374.60	1716	<.001	.029 (.028-.030)	-	.915	-	-
Débil	4930.20	1798	<.001	.026 (.025-.027)	.004	.925	.010	53.10
Fuerte	5182.90	1836	<.001	.026 (.025-.027)	.001	.926	.001	178.91***
Estricto	5389.90	1880	<.001	.026 (.025-.027)	.000	.922	.004	105.31***
Medias	6362.20	1886	<.001	.029 (.028-.030)	.018	.913	.009	166.80***

Fuente: elaboración propia *=p<.05 **=p<.01 ***=p<.001.

Recibido: 31/03/2023 Aceptado: 05/12/2023

los niveles de invariancia considerando los estudios de FP y el sexo autoinformado. El RMSEA mostró un ajuste adecuado con valores de .028 y .029 para los niveles de FP y el sexo autoinformado, respectivamente.

Antes de presentar los resultados del ajuste analítico de los parámetros del modelo bifactorial con 1 factor general y 5 factores específicos no correlacionados, se debe tener en cuenta que el supuesto general de este tipo de modelo es que el factor general explica una mayor cantidad de variancia de los ítems que los factores específicos. Sin embargo, si solo se consideran los valores de los índices de ajuste global, es posible concluir que el modelo bifactorial es el mejor, cuando de hecho no lo es. Esta posibilidad de aparición de falsos positivos se incrementa si se tiene en cuenta que los índices de ajuste tradicionales (RMSEA, CFI, TLI, etc.) tienden a favorecer los modelos bifactoriales (Gignac, 2016). Es necesario interpretar índices específicos para modelos bifactoriales (Omega H, índice H, PUC y ECV) para garantizar la plausibilidad de una verdadera estructura bifactorial con factores específicos relevantes, frente a la unidimensionalidad de un factor general. En este estudio, el Omega H fue de .72, el índice H fue de .88, el PUC fue de .80 y el ECV fue de .40 para el factor general de vinculación; estos valores indican un factor general sólido, pero no se puede asegurar la unidimensionalidad. Los valores de Omega H para los cinco factores específicos fueron todos superiores a .30 y los valores del índice H fueron todos superiores a .70, excepto para el primer factor específico Deberes escolares y disciplina (H = .61). Estos resultados indican que los cinco factores específicos pueden considerarse significativos en el modelo bifactorial (Tabla V).

En cuanto al valor de ECV-i, solo el 14% de los ítems (6 de 44) presentaron un valor igual o superior a .80, por lo que su variabilidad fue explicada casi en su totalidad por el factor general. Estos resultados reafirman el peso significativo de los factores específicos en la contribución a la variancia de los ítems explicados junto con el factor general. Los factores con mayor peso específico según el valor promedio del valor ECV-i de los ítems que los componen fueron, en este orden, Integración social (.20), Apoyo de la familia (.25) y Apoyo del profesorado (.39); mientras que Metas y expectativas (.63) y Deberes y disciplina escolar (.60) presentaron un menor peso específico en comparación con el factor general.

Finalmente, todas las cargas factoriales estandarizadas, tanto del factor general como de los específicos, fueron estadísticamente significativas ($p<.01$), variando de .24 a .56 para el factor general (M=.37) y de .06 a .68 (M=.44) para los factores específicos. No se han detectado ítems que

TABLA V. Cargas factoriales estandarizadas e índices de fiabilidad del modelo bifactorial 5F (44 ítems)

Ítems	Factor general de vinculación	Deberes y disciplina escolar (SD)	Metas y expectativas (GE)	Integración social (SI)	Apoyo del profesorado (TS)	Apoyo de la familia (FS)	ECV-i
SD1	.52	.51					.51
SD2	.44	.58					.36
SD3	.51	.24					.81
SD4	.46	.53					.44
SD5	.26	.24					.55
SD6	.24	.14					.75
SD7	.26	.17					.71
SD8	.33	.27					.61
SD9	.44	.30					.69
GE1	.38		.07				.97
GE2	.47		.06				.98
GE3	.48		.31				.70
GE4	.45		.23				.80
GE5	.42		.19				.84
GE6	.56		.45				.61
GE7	.53		.52				.51
GE8	.38		.44				.43
GE9	.44		.63				.32
GE10	.46		.51				.45
GE11	.32		.15				.81
GE12	.46		.51				.45
GE13	.44		.62				.33
SI1	.24			.68			.11
SI2	.29			.66			.16
SI3	.24			.54			.16
SI4	.30			.57			.21
SI5	.27			.36			.36
TS1	.35				.51		.29
TS2	.36				.56		.30
TS3	.30				.35		.42

(Continuada)

Recibido: 31/03/2023 Aceptado: 05/12/2023

TABLA V. Cargas factoriales estandarizadas e índices de fiabilidad del modelo bifactorial 5F (44 ítems) *(Continuada)*

Ítems	Factor general de vinculación	Deberes y disciplina escolar (SD)	Metas y expectativas (GE)	Integración social (SI)	Apoyo del profesorado (TS)	Apoyo de la familia (FS)	ECV-i
TS4	.33				.50		.30
TS5	.37				.52		.34
TS6	.37				.51		.34
TS7	.43				.35		.59
TS8	.35				.26		.65
TS9	.37				.58		.29
FS1	.28					.62	.17
FS2	.36					.62	.25
FS3	.32					.65	.20
FS4	.35					.42	.41
FS5	.35					.48	.35
FS6	.31					.65	.18
FS7	.33					.68	.19
FS8	.31					.60	.21
Omega H	.72	.32	.35	.35	.51	.67	
Índice H	.88	.61	.75	.73	.73	.82	
ECV-SS	.40	.45	.45	.82	.63	.77	
ECV-SG	.40	.07	.10	.12	.13	.17	

Fuente: elaboración propia. Todas las cargas son estadísticamente significativas a nivel de p<.001.

presenten simultáneamente cargas factoriales bajas en el factor general y en los factores específicos, por lo que no se ha eliminado ningún ítem de la estructura latente, también para preservar la validez de contenido del instrumento.

Discusión

La medicion de la vinculación del alumnado puede ser muy relevante para prevenir problemas escolares y evitar el fracaso escolar y, por tanto, el abandono escolar (Wang et al., 2011). Desafortunadamente, la

Recibido: 31/03/2023 Aceptado: 05/12/2023

literatura muestra un número limitado de instrumentos para evaluar dicha vinculación (Cedefop, 2016b; Christenson et al., 2012; Wang et al., 2019), lo que implica algunas limitaciones en cuanto al dominio de la medida y la población objetivo. Además, en el contexto educativo español no existen estudios dirigidos a medir la vinculación de los estudiantes en el ámbito de la educación secundaria profesional (Echeverría y Martínez, 2017).

El presente estudio proporciona un instrumento con puntuaciones fiables y válidas del constructo de vinculación estudiantil en una población estudiantil de FP, tanto de grado básico como medio. Los resultados mostraron cinco factores no correlacionados para operacionalizar el contructo de vinculación en esta población: deberes y disciplina escolar, metas y expectativas, integración social (iguales del grupo clase), apoyo del profesorado y apoyo de la familia. Estos hallazgos explican que el constructo muestra múltiples cargas de componentes sobre un factor general, la vinculación de los estudiantes. Este factor general refleja lo que se comparte entre los ítems y representa las diferencias de los estudiantes en la dimensión objetivo.

En este caso, el estudio sugiere que se pueden crear puntuaciones para diferenciar al alumnado con patrones de vinculación específicos (Cedefop, 2016a). Este resultado refuerza la idea de que la vinculación estudiantil no siempre implica el mismo grado de vulnerabilidad o desapego en todos los estudiantes. En esta línea, es esencial estudiar los factores implicados en el proceso. El contexto escolar, las experiencias personales vividas, la actividad escolar, el profesorado, el grupo de iguales y los progenitores, condicionan la vinculación del alumnado, de ahí la necesidad de medir diferentes factores de vinculación escolar.

Además, de acuerdo con los resultados, cinco ítems del apoyo de la familia no obtienen un buen comportamiento psicométrico. Estos ítems consideran una forma indirecta de medir el apoyo familiar, explicando la reacción de los progenitores ante ciertos comportamientos o sus expectativas, sin contemplar lo que hace la familia cuando el estudiante necesita ayuda o atención personal. Asimismo, este resultado podría explicar que el peso de estas variables familiares no sea el mismo para el alumnado de FPGB que para el de FPGM, de acuerdo con la etapa evolutiva (Alonso, 2014). Así, esta variable sería más importante entre el alumnado más joven, concretamente entre el de FPGB (Elffers, 2013), y

Recibido: 31/03/2023 Aceptado: 05/12/2023

no tan importante entre el de FPGM, que suele estar por encima de la edad teórica de ingreso (16 años).

Por su parte, ítems del factor dos, metas y expectativas, como "*Me comparo con mis compañeros/as para saber si voy aprendiendo al ritmo adecuado"* o *"Cuando hago una actividad escolar intento comprender bien lo que estoy haciendo"* pueden no corresponder a las instrucciones dadas por el profesorado, principalmente, en los estudios de FPGB y FPGM, justificando su mal comportamiento psicométrico. De acuerdo con la literatura (Fix et al., 2019), el profesorado de FP prioriza las instrucciones para evaluar el progreso intraindividual y proporcionar retroalimentación y apoyo social a los estudiantes en función de sus resultados.

La escasez de estudios empíricos sobre la vinculación escolar y el alumnado de educación secundaria profesional en España hace que la contribución de este estudio sea valiosa en sí misma dentro del contexto psicoeducativo y social. Además, el instrumento IVE con evidencia adecuada de fiabilidad y validez, es de gran utilidad para desarrollar un instrumento sencillo que permita detectar bajos niveles de vinculación estudiantil y también prevenir el abandono escolar que se produce en la mencionada población estudiantil. Asimismo, el instrumento IVE puede favorecer las intervenciones psicoeducativas preventivas y las adaptaciones escolares individualizadas, mejorando así la orientación educativa en estos temas (Pilcher et al., 2020).

Por otro lado, futuras investigaciones deberían incluir estudios longitudinales que permitan ampliar los resultados. Además, la administración del instrumento IVE podría complementarse con otras escalas que aborden el contexto escolar (organización escolar, metodología del centro, etc.) para completar la evaluación de la vinculación del alumnado y la prevención del abandono escolar.

Financiación

Los datos de este articulo provienen del proyecto I+D "Itinerarios de éxito y abandono en la formacion profesional de nivel I y II" (EDU2013-42854-R), financiado por el Ministerio de Economia y Competitividad, la Agencia Estatal de Investigacion y el Fondo Europeo de Desarrollo Regional (Gobierno de España).

Referencias Bibliográficas

Alonso, C. (2014). Familia, escuela y clase social: sobre los efectos perversos de la implicación familiar. [Family, school and social class: on the perverse effects of family involvement]. *Revista de Asociación de Sociología de la Educación, 7*(2), 395-409.

Appleton, J. J., & Christenson, S. L. (2004). *Scale description and references for the Student Engagement Instrument*. Unpublished manuscript.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427-445. https://doi.org/10.1016/j.jsp.2006.04.002

Betz, N., & Voyten, K. K. (1997). Efficacy and outcome expectation influence career exploration. *Career Development Quarterly*, *46*(2), 197-189. https://doi.org/10.1002/j.2161-0045.1997.tb01004.x

Brauer, K., Ranger, J., & Ziegler, M. (2023). Confirmatory factor analyses in psychological test adaptation and development. A nontechnical discussion of the WLSMV estimator. *Psychological Test Adaptation and Development, 4*, 4-12. https://doi.org/10.1027/2698-1866/a000034

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Publications.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.

Cedefop (2016a). *Leaving education early: Putting vocational education and training centre stage. Volume I: Investigating causes and extent*. Publications Office. Cedefop research paper, 57. https://doi.org/10.2801/893397

Cedefop (2016b). *Leaving education early: Putting vocational education and training centre stage. Volume II: Evaluating policy impact*. Publications Office. Cedefop research paper, 58. https://dx.doi.org/10.2801/967263

Cerdà-Navarro, A., Quintana-Murci, E., & Salvà-Mut, F. (2022). Reasons for dropping out of intermediate vocational education and training in Spain: the influence of sociodemographic characteristics and academic background. *Journal of Vocational Education & Training,* 1-25. https://doi.org/10.1080/13636820.2022.2049625

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. Springer. Recuperado a partir de https://www.researchgate.net/profile/Azkananda_Widiasani/publication/310773130_Handbook_of_Student_Engagement/links/5836a0dd08aed45931c772b7.pdf

Constantinou, M., & Fonagy, P. (2019). Evaluating Bifactor Models of Psychopathology Using Model-Based Reliability Indices. *PsyArXiv.* https://doi.org/10.31234/osf.io/6tf7j

Echeverría, B., & Martínez, P. (2017). *Diagnóstico de la investigación sobre la Formación Profesional Inicial en España (2005-2017)*. [Diagnosis of research on Initial Vocational Training in Spain (2005-2017)]. Fundación Bankia. Recuperado a partir de https://www.fundacionbertelsmann.org/wp-content/uploads/2021/10/2019_Publicacion_Diagnostico_Investigacion_FPI_digital.pdf

Elffers, L. (2013). Staying on track: Behavioral engagement of at-risk and non-at-risk students in post-secondary vocational education. *European Journal of Psychology of Education*, *28*(2), 545–562. https://doi.org/10.1007/s10212-012-0128-3

Elosua, P., & Zumbo, B. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema, 20*(4), 896-901. Recuperado a partir de https://reunido.uniovi.es/index.php/PST/article/view/8747

Estell, D. B., & Perdue, N. H. (2013). Social support and behavioral and affective school engagement:The effects of peers, parents, and teachers. *Psychology in the Schools*, *50*, 325-339. https://doi.org/10.1002/pits.21681

Fix, G. M., Ritzen, H. T. M., Pieters, J. M., & Kuiper, W. A. J. M. (2019). Effective curricula for at-risk students in vocational education: a study of teachers' practice. *Empirical Research in Vocational Education and Training*, *11*, 1-17. https://doi.org/10.1186/s40461-018-0076-5

Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology*, *43*(3), 338-346. https://doi.org/10.1037/0022-0167.43.3.338

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59-109. https://doi.org/10.3102/00346543074001059

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence, 55*, 57-68. https://doi.org/10.1016/j.intell.2016.01.006

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2008). *Multivariate Analysis* (5.ª Ed.). Prentice Hall.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hudson, G. (2020). *EGAnet: Exploratory Graph Analysis–A framework for estimating the number of dimensions in multivariate data using network psychometrics*. R package version 0.9.5.

Işik, E. (2013). Perceived Social Support and Locus of Control as the predictors of Vocational Outcome Expectations. *Educational Sciences: Theory & Practice*, *13*(3), 1426-1430. Recuperado a partir de https://files.eric.ed.gov/fulltext/EJ1017701.pdf

Janosz, M., Archambault, I., Lacroix, M., & Lévesque, J. (2007). *Trousse d'évaluation des décrocheurs potentiels (TEDP): Manuel d'utilisation*. Montréal: Groupe de recherche sur les environnements scolaires. Université de Montréal.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling. R package version 0.5-5*. Recuperado a partir de https://CRAN.R-project.org/package=semTools

Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling*. Guilford Press.

Le Blanc, M. (1996). MASPAQ, mesures de l'adaptation sociale et personnelle pour les adolescents québécois: manuel. Montréal, École de psycho-éducation, Groupe de recherche sur l'inadaptation psychosociale chez l'enfant, Université de Montréal.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*, 79-122. https://doi.org/10.1006/jvbe.1994.1027

Mameli C., & Passini S. (2017). Measuring four-dimensional engagement in school: A validation of the student engagement scale and of the agentic engagement scale. *TPM - Testing, Psychometrics, Methodology*

in Applied Psychology, 24(4), 527-541. https://doi.org/10.4473/TPM24.4.4

Martínez-Abad, F., & Rodríguez-Conde, M. J. (2017). Behaviour of product-moment and tetrachoric-polychoric correlations in ordinal scales: a simulation study. *RELIEVE, 23*(2), 1-19. doi: http://dx.doi.org/10.7203/relieve.23.2.9476

Massara, G. P., Di Matteo, T., & Aste, T. (2017). Network filtering for big data: Triangulated maximally filtered graph. *Journal of Complex Networks, 5*(2), 161-178. https://doi.org/10.1093/comnet/cnw015

McWhirter, E. H., Crothers, M., & Rasheed, S. (2000). The effects of high school career education on social-cognitive variables. *Journal of Counseling Psychology*, *47*(3), 330-341. https://doi.org/10.1037//0022-0167.47.3.330

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Technical report.* Recuperado a partir de https://www.statmodel.com/download/Article_075.pdf

Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical Analysis with Latent Variables: User's Guide* (Version 8). Muthén & Muthén.

Niittylahti, S., Annala, J., & Mäkinen, M. (2019). Student engagement at the beginning of vocational studies. *Nordic Journal of Vocational Education & Training, 9*(1), 21-42. https://doi.org/10.3384/njvet.2242-458X.199121

OECD (2023). *Propuestas para un plan de acción para reducir el abandono escolar temprano en España*. OECD Publishing. https://doi.org/10.1787/9bc3285d-es

Olmos Rueda, P., Mas Torello, O., & Salva Mut, F. (2020). Educational disengagement profiles: a multidimensional contribution within basic vocational education and training. *Revista de Educación*, (389), 67-91. https://doi.org/10.4438/1988-592X-RE-2020-389-455

Pilcher, S.; Torii, K., & Andow, Jo. (2020). *Evidence scan of educational interventions for children and young people disengaged from education*. Mitschell Institute.

Poli, D., & Hungler, B. (2000). *Investigación científica de la Salud*. [Scientific Health Research]. McGraw-Hill.

Quin, D., Heerde, J. A., & Toumbourou, J. W. (2018). Teacher support within an ecological model of adolescent development: Predictors of school engagement. *Journal of School Psychology*, *69*, 1-15. https://doi.org/10.1016/j.jsp.2018.04.003

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Recuperado a partir de https://www.R-project.org/

Reschly, A. L., & Christenson, S. L. (2012). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct. In S.L. Christentson, A. Reschly, & C. Wylie (Ed.) *Handbook of Research on Student Engagement* (pp 3-20). Springer.

Reise, S.P. Scheines, R., Widaman, K.F., & Haviland, M.G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement, 73*(1), 5-26. https://doi.org/10.1177/0013164412449831

Rosseel, Y. (2012). "lavaan": An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Soler, A., Martínez Pastor, J.I., López-Messeguer, R., Valdés, M.T., Sancho Gargallo, M.A., Morillo, B., & de Cendra, L. (2021). *Mapa del abandono educativo temprano en España. Informe General*. Fundación Europea Educación y Sociedad.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, *42*(5), 893-898. https://doi.org/10.1016/j.paid.2006.09.017

Stucky, B. D., & Edelen, M. O. (2015). Using heierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment*, (pp. 183-206). New York: Routledge.

Sureda, I., Jiménez, R., Álvarez, O., & Quintana, E. (2021). Emotional and Behavioural Engagement among spanish Student in vocational education training. *Sustainability*, *13*(7), 3882. https://doi.org/10.3390/su13073882

Wang, M. T., Fredricks, J., Ye, F., Hofkens, T., & Linn, J. S. (2019). Conceptualization and assessment of adolescents' engagement and disengagement in school: A Multidimensional School Engagement Scale. *European Journal Psychological Assess*ment, *35(*4), 592-606. https://doi.org/10.1027/1015-5759/a000431

Recibido: 31/03/2023 Aceptado: 05/12/2023

Wang, M. T., Willett, J. B., Eccles, J. S. (2011). The assessment of school engagement: examining dimensionality and measurement invariance by gender and race/ethnicity. *Journal of School Psychology*, *49*(4), 465-480. https://doi.org/10.1016/j.jsp.2011.04.001

Información de contacto: Rafael Jiménez. Universitat de les Illes Balears. Cra. de Valldemossa, km 7.5, 07122. Palma, Islas Baleares, España.
E-mail: rafa.jimenez@uib.es

Recibido: 31/03/2023 Aceptado: 05/12/2023

Anexo

Contenido de los ítems y sus estadísticas descriptivas por factores del modelo latente bifactorial 5F del IVE. Estadística descriptiva: [*N* válido, *M*, *SD,* asimetría (*g1*), apuntamiento (*g2*)]

Factor 1: Deberes y disciplina escolar (SD) (9 ítems)
SD1 Me esfuerzo con las tareas escolares [4509, 2.1, 0.7, 0.43, -0.54]
SD2 Estudio y/o hago deberes casi todos los días [4494, 1.78, 0.86, -0.64, -0.23]
SD3 Cuando hago un trabajo en el instituto/colegio, quiero hacerlo bien [4485, 2.52, 0.59, 1.48, -1.08]
SD4 Dedico tiempo suficiente fuera del instituto/colegio para hacer los deberes y estudiar [4490, 1.67, 0.84, -0.55, -0.17]
SD5* Molesto en clase a propósito [4501, 0.38, 0.66, 3.57, 1.9]
SD6* Contesto al profesor/a de forma poco educada [4489, 0.44, 0.79, 2.57, 1.82]
SD7* Utilizo "chuletas" u otros medios para copiar en un examen [4491, 0.42, 0.75, 2.57, 1.81]
SD8* He faltado a clase/s sin justificación [4486, 0.83, 0.95, -0.55, 0.79]
SD9 Antes de entregar mis trabajos o tareas académicas, las reviso para comprobar que las he hecho correctamente [4496, 2.13, 0.76, 0.36, -0.71]
Items eliminados del Factor 1:
1. Me gusta ir al instituto/colegio [4500, 1.83, 0.83, -0.04, -0.58]

Factor 2: Metas y expectativas (GE) (13 ítems)
GE1 Cuando me esfuerzo en los estudios, los resultados que obtengo son positivos [4467, 2.27, 0.73, 0.22, -0.75]
GE2 Considero que los exámenes, pruebas o ejercicios de clase son una buena herramienta para saber lo que he aprendido [4483, 2.06, 0.8, 0.3, -0.74]
GE3 Lo que estoy aprendiendo en las clases es importante para mi futuro profesional [4477, 2.45, 0.7, 1.34, -1.22]
GE4 Estudiar me va a ofrecer muchas oportunidades profesionales de futuro [4481, 2.49, 0.65, 1.19, -1.15]
GE5 Deseo seguir formándome una vez acabe los estudios actuales [4474, 2.3, 0.78, 0.58, -0.99]
GE6 Los estudios que estoy realizando me hacen ser optimista en relación con mi futuro profesional [4479, 2.24, 0.73, 0.62, -0.82]
GE7 Estudio porque me gusta lo que estoy haciendo [4482, 2.24, 0.82, 0.45, -0.97]
GE8 Gracias a los estudios que realizo podré tener un trabajo con el que ganaré suficiente dinero para vivir [4484, 2.1, 0.74, 0.46, -0.66]
GE9 Gracias a los estudios que realizo creo que podré dedicarme a lo que quiero [4480, 2.09, 0.82, 0, -0.68]
GE10 Mis estudios me ayudarán a tener éxito en mi carrera profesional [4470, 2.19, 0.69, 0.79, -0.7]
GE11 Para dedicarme a lo que realmente quiero, tendré que seguir formándome después de acabar estos estudios [4466, 2.31, 0.79, 0.4, -0.98]
GE12 Los estudios que realizo son adecuados a mis características personales [4481, 2.11, 0.74, 0.51, -0.67]
GE13 Me gusta la profesión para la que me estoy formando [4492, 2.25, 0.79, 0.66, -0.98]
Items eliminados del Factor 2:
1. Cuando hago una actividad escolar intento comprender lo que estoy haciendo [4489, 2.3, 0.61, 0.65, -0.51]
2. Me comparo con mis compañeros/as para saber si voy aprendiendo al ritmo adecuado [4482, 1.6, 0.93, -0.8, -0.19]

Recibido: 31/03/2023 Aceptado: 05/12/2023

Factor 3: Integración social (SI) (5 ítems)
SI1 Mis compañeros/as se preocupan por mí [4500, 1.91, 0.72, 0.55, -0.57]
SI2 Mis compañeros/as me ayudan cuando lo necesito [4505, 2.15, 0.67, 0.88, -0.61]
SI3 Mis compañeros/as respetan mis opiniones [4481, 1.96, 0.69, 0.7, -0.54]
SI4 Me gusta comunicarme con mis compañeros/as [4486, 2.37, 0.65, 1.14, -0.88]
SI5 Tengo amigos/as en mi instituto/colegio [4493, 2.49, 0.66, 1.85, -1.28]

Factor 4: Apoyo del profesorado (TS) (9 ítems)
TS1 Mis profesores/as están disponibles cuando los/las necesito [4509, 2.16, 0.64, 0.86, -0.5]
TS2 El profesorado de mi centro educativo escucha al alumnado [4503, 2.1, 0.64, 0.7, -0.44]
TS3 Las normas del centro educativo donde estudio son justas [4474, 1.9, 0.76, 0.1, -0.47]
TS4 El profesorado de mi centro educativo está interesado en mí como persona, no sólo como estudiante [4462, 1.79, 0.79, -0.16, -0.39]
TS5 En general, mis profesores/as son abiertos/as y honestos/as conmigo [4475, 2.12, 0.67, 0.52, -0.48]
TS6 En general, el profesorado de mi instituto/colegio trata de manera adecuada al alumnado [4475, 2.15, 0.68, 0.59, -0.57]
TS7 Me gusta hablar con los/las profesores/as de mi instituto/colegio [4469, 1.88, 0.76, 0.14, -0.5]
TS8 Me siento seguro/a en mi instituto/colegio [4473, 2.19, 0.69, 0.93, -0.72]
TS9 En mi instituto/colegio, la mayoría del profesorado se preocupa por el alumnado [4484, 1.98, 0.71, 0.37, -0.49]

Factor 5: Apoyo de la familia (FS) (8 ítems)
FS1 Mi familia está disponible cuando la necesito [4498, 2.49, 0.75, 1.6, -1.45]
FS2 Cuando sucede algo bueno en el instituto/colegio, mi familia quiere saberlo [4492, 2.33, 0.78, 0.77, -1.08]
FS3 Cuando tengo problemas en el instituto/colegio, mi familia está dispuesta a ayudarme [4463, 2.47, 0.73, 1.88, -1.44]
FS4 Mis padres** saben cuándo tengo deberes o exámenes [4354, 1.56, 1, -1.05, -0.11]
FS5 Si tengo un problema en el instituto/colegio, normalmente lo hablo con mis padres** [4396, 1.71, 1, -0.97, -0.3]
FS6 Mis padres** hacen todo lo que pueden para ayudarme a tener buenas notas [4401, 2, 0.88, -0.26, -0.64]
FS7 Puedo contar con mis padres** cuando tengo dificultades en el instituo/colegio [4396, 2.18, 0.85, 0.24, -0.91]
FS8 Mis padres** me preguntan habitualmente cómo me van las cosas en el instituto/colegio [4406, 2.3, 0.81, 0.61, -1.05]
Items eliminados del Factor 5:
1. Mi familia quiere que intente resolver por mí mismo los problemas que puedo tener en el instituto/colegio [4320, 2.15, 0.78, 0.3, -0.75]
2. Mis padres** esperan que continúe mis estudios durante el mayor tiempo posible [4399, 2.49, 0.69, 1.77, -1.37]
3. Para mis padres** es importante que yo apruebe el curso [4362, 2.71, 0.56, 5.21, -2.14]
4. Daría un disgusto a mis padres** si dejara el instituto/colegio [4360, 2.35, 0.89, 0.71, -1.27]
5. Para mis padres** es importante que dé lo mejor de mí en el instituto/colegio [4370, 2.61, 0.6, 2.66, -1.56]

** Items invertidos ** En la encuesta se utiliza la expresión "padres (o personas que cumplen sus funciones)"*

Recibido: 31/03/2023 Aceptado: 05/12/2023

¿Quién comparte fake news? Consumo y distribución de información entre adolescentes y su relación con el discurso de odio

Who shares fake news? The consumption and distribution of information among adolescents and its relationship to hate speech

https://doi.org/10.4438/1988-592X-RE-2024-406-645

María Reneses
https://orcid.org/0000-0002-9708-6896
Instituto de Investigación Tecnológica, Escuela Técnica Superior de Ingeniería ICAI, Universidad Pontificia Comillas
Centro de Ciencias Humanas y Sociales. Universidad Pontificia Comillas

María Riberas-Gutiérrez
https://orcid.org/0000-0003-2030-0310
Centro Universitario Cardenal Cisneros

Nereida Bueno-Guerra
https://orcid.org/0000-0003-1442-7905
Centro de Ciencias Humanas y Sociales. Universidad Pontificia Comillas

Resumen

Los jóvenes son especialmente vulnerables a las fake news, ya que tienden a informarse principalmente a través de las redes sociales. Además, gran parte del contenido de estas noticias falsas incluye mensajes de odio y, por tanto, persigue la discriminación contra las minorías. Este artículo presenta los resultados de ocho grupos de discusión europeos y una encuesta representativa a nivel regional (Comunidad de Madrid) sobre cómo los jóvenes acceden y comparten la información online, incluyendo las fake news. El estudio forma parte del proyecto

Recibido: 31/03/2023 Aceptado: 22/12/2023

europeo H2020 RAYUELA sobre menores y riesgos de Internet. Los resultados muestran una preferencia generalizada entre los jóvenes por informarse a través de las redes sociales, a pesar de admitir que los medios tradicionales son más fiables. Esta predilección se explica porque las primeras utilizan más vídeos, congregan a más gente y publican comentarios con opiniones contrarias. También se describen las diferencias entre los adolescentes que comparten noticias falsas a propósito y sin querer. Los segundos afirman haber empleado en mayor medida dispositivos manuales (por ejemplo, smartphone, tablet), mientras que los primeros admiten en mayor proporción haber sido también autores de discursos de odio en línea (sexismo, racismo y LGTBIfobia).

Palabras clave: fake news, internet, discurso del odio, información online, adolescentes, ciberodio.

Abstract

Young people might be especially vulnerable to fake news as they tend to get informed mainly online. Besides, much of the content of these fake news may contain hate speech and therefore support discrimination against minorities. This article presents the results of eight European focus groups and a regionally (Madrid) representative survey about how young people access and share fake news information online carried out in the framework of European Project H2020 RAYUELA. The results show a general preference among young people to get information from social networks despite admitting that traditional media are more reliable because the former use more videos, congregate more people and post comments with opposing opinions. We also describe differences between teenagers sharing fake news on purpose or unintentionally. The latter tend to more frequently use hand devices (i.e., smartphone, tablet), while the former more widely admit having become offenders of online hate speech (sexism, racism and LGTBIphobia).

Keywords: fake news, internet, hate speech, online information, teenagers, cyberhate.

Introducción

La difusión de noticias falsas (fake news) ha aumentado en los últimos años, especialmente en el último periodo de crisis sanitaria (Apuke y Omar, 2020). Pese a que no existe una clara definición del término (Anderau, 2021) para este artículo las entendemos como contenidos elaborados intencionalmente para engañar al lector haciéndole creer

Recibido: 31/03/2023 Aceptado: 22/12/2023

que la información que está recibiendo es verdadera (Shu et al., 2017). Precisar la intención sirve para distinguirlas, por ejemplo, de la sátira (Verstrade et al., 2018). Normalmente se distribuyen a través de las webs y redes sociales, pero en ocasiones también se producen a través de los principales medios de comunicación tradicionales (Muigai, 2017).

Uno de los principales peligros de las fake news es que se difunden con mayor rapidez, profundidad y amplitud que las noticias verdaderas (Vziatysheva, 2020). Los mismos elementos que explican que una noticia se considere atractiva (su contenido emocional, su contenido llamativo que sitúa al que la comparte como bien informado, o la potencial relevancia para el receptor), son también características de las noticias falsas. Además, en los entornos online, por el efecto del "filtro burbuja" y los algoritmos de las redes sociales, se acaban generando "cámaras de eco", por las que determinados contenidos se convierten en dominantes para el usuario a través de la repetición, dejando fuera otros puntos de vista alternativos que podrían enfrentarse a dichos contenidos (Fernández-García, 2017).

Esta rápida propagación plantea la cuestión sobre la motivación última de compartir material falso. Duffy y colaboradores (2020) descubrieron que, en muchos casos, las personas comparten noticias falsas sin saber que lo hacen bajo la motivación de ayudar o informar a otros. De hecho, detrás de la difusión de desinformación en las redes sociales hay un fuerte componente relacional: si el remitente original es un amigo, el vínculo de confianza previo establecido reduce el nivel de sospecha y, posteriormente, la necesidad de verificar la información. Dado que los jóvenes acceden a la información principalmente a través de las redes sociales (Herrero-Diz et al., 2020), el contenido procedente de amigos sería más valorado que el procedente de la fuente original (Herrero-Diz et al., 2021): se sabe que la relación de confianza con la persona que comparte la información afecta a la credibilidad del contenido (Tompson, 2017).

Es importante destacar que la autoexpresión y la socialización también podrían ser razones para que las personas compartan información falsa, dándole mayor importancia que a la exactitud de la información o la autoridad de la fuente (Chen et al., 2015). De hecho, el interés por compartir noticias o información podría prevalecer incluso si los sujetos saben que el contenido no es veraz (Herrero-Diz et al., 2020). Así, compartir o no determinados contenidos puede estar motivado también por la necesidad de mostrar la propia identidad o afinidad con determinadas ideas (Marwick, 2018).

Recibido: 31/03/2023 Aceptado: 22/12/2023

En este sentido, los medios tradicionales no siempre conectan con los intereses de los jóvenes (Marchi, 2012), pueden no ser percibidos necesariamente como más fiables (Chen et al., 2015) y utilizan de forma menos frecuente los formatos en los que más confían los adolescentes, como son los vídeos (Literat et al., 2020). Mendiguren y sus colegas (2020) describen la Generación "Z" cómo caracterizada por la interacción, la rapidez y la impaciencia, aspectos que facilitan caer en la desinformación. De hecho, los autores descubrieron que, aunque la mayoría de los jóvenes (82%) reconocía haber asumido como cierta alguna información falsa, les preocupaba la credibilidad de la información consultada. Sin embargo, parece que los jóvenes son fácilmente engañados si la información falsa se presenta con la forma y el formato utilizados por las noticias (Herrero-Diz et al., 2020), así como si el contenido está patrocinado y aporta datos y estadísticas (McGrew et al., 2017), ya que la apariencia desinteresada o científica se interpreta como sinónimo de credibilidad.

Además, las fake news podrían vincularse con el discurso del odio, un fenómeno que también ha ido en aumento en los últimos años (Mascheroni & Ólafsson, 2014; Hawdon et al., 2015), con consecuencias que trascienden el mundo virtual (Williams et al., 2020). Concretamente, los bulos raciales tienen un efecto que valida actitudes discriminatorias y racistas (Cerase & Santoro, 2018), incrementando el rechazo contra migrantes y refugiados (Schäfer & Schadauer, 2018). En cuanto al sexismo, hay pruebas de que, por ejemplo, los estereotipos de género se refuerzan a través de las noticias falsas cuando se muestra a las candidatas electorales como no aptas para el liderazgo, ya sea al villanizarlas o al trivializarlas (Stabile et al., 2019).

Así pues, parece más relevante que nunca educar a la población en el uso adecuado de las tecnologías que incluyen difusión de información. Hacerlo implica considerar que la ciudadanía mediática puede ser empoderada a través de competencias comunicativas que las protejan frente a esos riesgos (Gozálvez y Contreras-Pulido, 2014) y les permitan un correcto desarrollo de su libertad de expresión. A esta preocupación por la alfabetización mediática, ya identificada por la ONU (Aguaded, 2012) y la UNESCO (2009) durante las últimas décadas, y expresada de manera recurrente a través de programas específicos dirigidos a prevenir las conductas de riesgo o extremismo (e.j., Al Nasser, 2017), se la conoce como educomunicación. Este paradigma entiende como necesaria la educación sobre la manera de recibir y emitir comunicación basada en la

Recibido: 31/03/2023 Aceptado: 22/12/2023

participación y diálogo de sus agentes activos, en oposición a los modelos educativos transmisivos (Aparici, 2011). Bajo este paradigma, resulta entonces relevante conocer la manera de relacionarse de los agentes más vulnerables en el uso de las tecnologías, como son los adolescentes, así como los medios que emplean y sus motivaciones, de manera que se puedan proponer vías preventivas de educación basadas en la evidencia.

Con fines preventivos, resulta relevante comprender, en primer lugar, su propia experiencia en Internet, es decir, las formas en que consumen y difunden información. En segundo lugar, las principales razones y los factores de riesgo potenciales que están detrás de la difusión de noticias falsas. Para ello, en este estudio exploratorio abordamos el fenómeno de las fake news a través de ocho grupos de discusión de adolescentes en cinco países europeos diferentes (Portugal, España, Grecia, Eslovaquia y Estonia) y después realizamos una encuesta representativa en uno de ellos (España). Nuestra justificación para seguir esta metodología es doble: siendo las fake news un fenómeno de estudio reciente, no existe mucha evidencia sobre las tendencias y motivaciones de los jóvenes a la hora de compartir noticias falsas, ni sobre si existen algunas diferencias entre quienes comparten información falsa de forma intencionada y sin saberlo. Con un estudio exploratorio, podemos contribuir con información preliminar sobre los principales aspectos implicados, y, escuchando sus experiencias en primera persona. Por ello, en los grupos de discusión internacionales recogimos primero datos generales y posibles tendencias para luego triangular la información obtenida a nivel local a través de una encuesta regional representativa en Madrid.

En términos de prevención, explorar las fuentes que los participantes utilizan para informarse y las características en las que se basan para juzgar la credibilidad del contenido es útil para priorizar en los programas aprender a comprobar la veracidad de las fuentes o hacer hincapié en la relevancia de consumir información veraz y en los peligros de confiar en información falsa. Además, queremos explorar dos formas diferentes de difundir información falsa (a propósito, y de forma involuntaria): las diferentes motivaciones de una y otra forma de compartir información falsa podrían tenerse en cuenta en los programas preventivos. Por último, al explorar la prevalencia y la experiencia en primera persona de la incitación al odio online, podemos tener una idea preliminar sobre cómo perciben los adolescentes este fenómeno y si existen algunas características que deberían detectarse o abordarse para prevenir la agresión y la victimización.

Método

En este estudio seguimos dos metodologías. En primer lugar, realizamos un análisis cualitativo exploratorio mediante un análisis de contenido categórico de ocho grupos de discusión organizados en cinco países europeos diferentes (Portugal, España, Grecia, Eslovaquia y Estonia). En segundo lugar, realizamos un análisis cuantitativo mediante una encuesta representativa en la Comunidad de Madrid (España). Ambos procesos fueron aprobados por el Comité de Ética de la universidad de los autores antes de iniciar la recogida de datos. En ambos casos (grupos de discusión y encuesta) se explicó a los estudiantes el proyecto y la voluntariedad de su participación, y los padres o los participantes, según su edad, firmaron los consentimientos informados. Aunque las transcripciones del grupo de discusión no se comparten en el material suplementario por motivos de confidencialidad, los datos brutos de la encuesta pueden descargarse de nuestro repositorio institucional (link: http://hdl.handle.net/11531/69582).

Análisis cualitativo

A partir de la revisión bibliográfica, diseñamos las preguntas de los grupos de discusión con el objetivo de conocer la opinión de los participantes sobre su forma de consumir y de compartir noticias, sus criterios de verificación y su experiencia directa con las fake news (véase el anexo I en el material suplementario para el conjunto de preguntas iniciales).

Realizamos ocho grupos de discusión (Tabla I) compuestos por un total de 47 adolescentes de 12-14 o 14-17 años en cinco países europeos (Portugal, España, Grecia, Eslovaquia y Estonia). Aunque la selección de los candidatos se hizo al azar, se buscó el equilibrio entre chicos y chicas y, cuando fue posible, de diversidad (etnia y orientación sexual). Los entrevistadores fueron investigadores formados con experiencia de trabajo de campo en victimización infantil en línea.

Las intervenciones en los grupos de discusión se transcribieron íntegramente y se tradujeron al inglés -ya que cada grupo se realizó en su lengua local-, y se analizaron mediante un análisis de contenido categorial, es decir, se descompuso el texto en unidades que se agruparon en categorías según áreas temáticas (Páramo, 2011). Finalmente, a

Recibido: 31/03/2023 Aceptado: 22/12/2023

TABLA I. Características de los grupos de discusión

País	Total de grupos	Número de participantes por grupo	Edad por grupo
España	2	6 participantes (SP1) 5 participantes (SP2)	12-14 14-17
Portugal	1	5 participantes (P1)	14-17
Grecia	1	6 participantes (G1)	14-16
Estonia	3	6 participantes (E1) 5 participantes (E2) 7 participantes (E3)	12-14 12-14 14-17
Eslovaquia	1	7 participantes (SL1)	14-16

Fuente: Elaboración propia.

partir de los datos extraídos con el sistema de categorías, realizamos la conceptualización e interpretación. Siguiendo a Gil y colaboradores (1994), en primer lugar, nos acercamos a los datos leyendo el discurso completo y anotando algunas ideas clave. En segundo lugar, realizamos un análisis descriptivo, estructurando toda la información en torno a los temas principales. Dentro de esta estructura, incluimos citas para ilustrar las ideas más relevantes. En tercer lugar, realizamos un análisis de reducción en dos etapas: segmentación, dividiendo el texto en unidades, y categorización, agrupando las unidades conceptualmente. Las categorías se extrajeron directamente del guion inicial del debate e inductivamente de los temas planteados durante los grupos de discusión. Tras la reducción de datos, trazamos y creamos una matriz textual con los grupos y las categorías. Para ello, buscamos significados globales y relacionamos ideas similares. Este análisis fue realizado en todas las fases por al menos dos investigadores.

Análisis cuantitativo

Nuestro objetivo era contrastar localmente algunas de las conclusiones de los grupos de discusión. Para ello, distribuimos una encuesta en una muestra representativa de la Comunidad de Madrid (España) compuesta por 682 participantes (nivel de confianza de estimación del 95%). Diseñamos la encuesta a partir de los resultados de los grupos de

discusión y la trasladamos a la plataforma de Microsoft. La encuesta incluyó preguntas sociodemográficas sobre la edad, el sexo, el origen migratorio y la orientación sexual, información sobre el uso de Internet, incluido el tiempo que pasan conectados y las aplicaciones y dispositivos que utilizan con más frecuencia, y preguntas relacionadas con la prevalencia de la difusión de noticias falsas y con la importancia otorgada a distintos aspectos a la hora de compartir contenido online. El conjunto completo de preguntas figura en el Anexo 2 del material suplementario.

Para seleccionar a los participantes utilizamos un método de muestreo probabilístico estratificado. Los estratos estaban formados por el tipo de colegio al que pertenecían los participantes (público, privado y concertado) y el tipo de lugar en el que vivían (gran ciudad, ciudad media y pueblo). Los participantes tenían entre 13 y 17 años de edad para cubrir el rango de edades consideradas en los grupos de discusión (véase la tabla II).

La encuesta se realizó en línea a través de Microsoft Forms y se contestó durante una clase en los centros escolares. Incluía preguntas sobre los hábitos de los jóvenes en Internet, y preguntas sobre las noticias falsas y el discurso de odio online. Las preguntas se diseñaron para contrastar algunas de las preguntas e hipótesis de investigación que surgieron de los grupos de discusión: ¿Hasta qué punto están extendidas las noticias falsas?; ¿Existe un perfil del que comparte noticias falsas?; ¿Este perfil está relacionado con el grado de importancia que se da a los distintos temas a la hora de compartir información?; ¿Cómo se relaciona con convertirse en agresor de discurso de odio?

TABLA II. Características de la muestra de la encuesta

Colegios	n	Porcentaje del total	Tipo de centro educativo	Tipo de ciudad[1]
A	109	16.0%	Privado	Grande
B	64	9.4%	Privado	Grande
C	107	15.7%	Concertado	Mediana
D	70	10.3%	Público	Mediana
E	140	20.5%	Público	Grande
F	192	28.2%	Público	Pequeña

Fuente: Elaboración propia. Nota: Se consideró ciudades "grandes" a las que tenían más de 100.000 habitantes; las "medianas", entre 10.000 y 100.000, y las "pequeñas", <10.000.

Recibido: 31/03/2023 Aceptado: 22/12/2023

Para la explotación estadística, primero se revisó y preparó la base de datos mediante depuración (es decir, supresión de casos duplicados, casos atípicos), realización de un análisis de coherencia y codificación. Algunas variables se recodificaron para facilitar el análisis posterior, y los datos se exportaron y analizaron con el programa estadístico IBM/SPSS versión 28. Además del análisis de frecuencias simples, se realizó un análisis bivariante mediante tablas de contingencia y la prueba de Chi-cuadrado, verificando la significación estadística entre pares de variables mediante los residuos tipificados corregidos y aportando el coeficiente de contingencia (CC).

Resultados

Fuentes de información y fiabilidad de los medios de comunicación

Para la mayoría de los participantes en los grupos de discusión, las redes sociales, y no los medios de comunicación tradicionales, aparecen como la principal fuente de información. No obstante, hubo una pequeña diferencia entre las edades. En los grupos de discusión más jóvenes (12-14 años), salvo una participante que veía las noticias en la televisión con sus padres, los medios tradicionales no fueron mencionados como consultados, ni siquiera en formato online. Aunque algunos participantes refirieron utilizar Google para buscar información, declararon mayoritariamente ver Tik Tok y YouTube, mostrando preferencia por el formato de vídeo. En concreto, se expresó que YouTube se utiliza para "*buscar algo visto anteriormente en Internet en general o en la televisión, para comprobar qué información da YouTube*", y también para "*un repaso rápido de las noticias*" (E1). En los grupos de mayor edad (14-17), hubo una mayor diversidad. Aunque algunos refirieron acudir a los medios tradicionales para informarse, la fuente señalada por la mayoría fueron las redes sociales, principalmente Tik Tok y Facebook:

"*Tengo notificaciones en Facebook cuando salen nuevas noticias*" (E3). Esto significa que la mayoría de las noticias no son seleccionadas por los espectadores, sino que aparecen de acuerdo con los algoritmos de preferencia del usuario a través de las notificaciones: "*No es tanto como que te aparezcan noticias como que por las noticias que sube la gente o las cosas que cuenta la gente te vas enterando*" (SP2). De hecho,

algunos participantes reconocieron que las redes sociales "*no son una buena fuente de información porque* [sólo] *se ven cosas que siguen las propias preferencias*" (P1).

A pesar del uso preferente de las redes sociales, la mayoría de los participantes en los grupos de discusión se refirieron a los medios tradicionales como la fuente de información más creíble, concretamente a las noticias en televisión, explicando que "*cuando algo aparece en las noticias se siente más seguro*" (SP2). Sin embargo, su falta de confianza general en las redes sociales no afecta a su patrón de acceso a la información: la mayoría se informa a través de las redes sociales porque allí "*aparecen primero las noticias de última hora"* (P1), "*está todo el mundo"* (G1) y "*puedes seguir las polémicas*" (G1). Sólo hubo unos pocos casos en los que algunos jóvenes refirieron informarse a través de medios tradicionales (por ejemplo, tres participantes veían las noticias en la televisión con sus padres, un joven de los grupos de más edad leía el periódico online y algunos participantes accedían a la BBC y la aplicación de noticias locales). Ningún participante afirmó leer el periódico en papel, y muchos de ellos reconocieron que rara vez leían noticias ("*No leemos mucho las noticias*" E3).

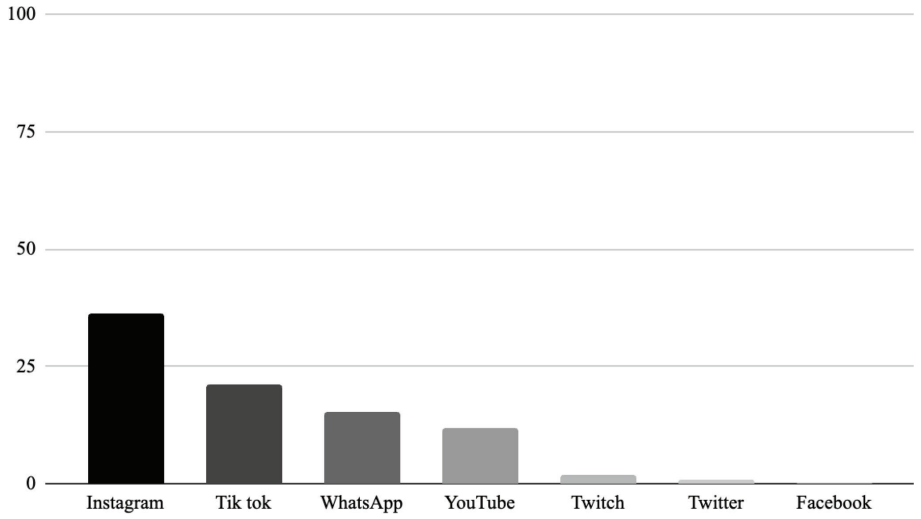
La encuesta no abordó los diferentes canales de información, pero sí midió las redes sociales más utilizadas, que se resumen en la figura I. En el momento en que se realizó la encuesta (diciembre de 2021), Instagram, Tik Tok, Whatsapp y Youtube eran las aplicaciones más utilizadas.

Precauciones y preocupaciones de los adolescentes al compartir información en Internet

Aunque la fiabilidad no es una de las razones por las que los participantes en los grupos de discusión eligen una determinada plataforma para informarse, eso no significa que carezcan de criterios para valorar qué información es fiable y cuál no. En general, los participantes señalaron como lo más relevante que el emisor fuera de confianza y que el sitio de donde procede la información también lo fuera, describiendo que hay "*páginas famosas*" como los grandes periódicos o Wikipedia que les parecen fiables. Por último, en menor medida, la autoría también se consideró un factor de credibilidad, ya fuera por ser un autor reconocido o un usuario verificado, por ejemplo, de YouTube.

Recibido: 31/03/2023 Aceptado: 22/12/2023

FIGURA I. Porcentaje de aplicaciones más utilizadas (N=682)



Fuente: Elaboración propia.

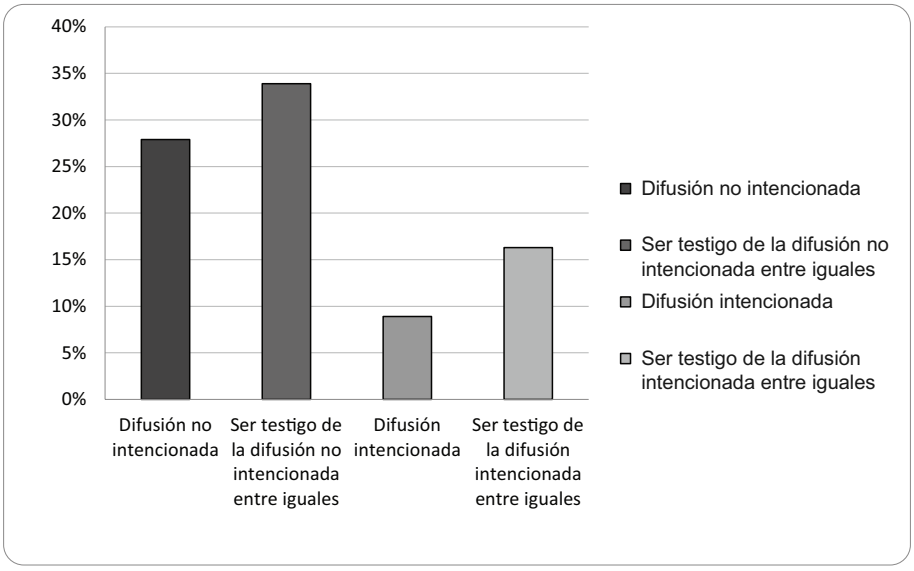
Otras cuestiones que también se sugirieron fueron el estilo (es decir, buena saturación de los colores, fotografías de buena calidad y ausencia de erratas, titulares poco acertados o mala ortografía), los comentarios de la publicación (por ejemplo, "*Fíjate ante todo en los comentarios. Si los comentarios dicen que es falso, no te fíes*", SP2), la fecha de publicación y la inclusión de un vídeo. La mayoría de los participantes señalaron que apenas contrastaban la información ("*sólo si es muy incompleta*" (P1)). Algunos refirieron que sólo contrastaban en caso de querer compartir, prefiriendo a veces no compartir que contrastar: "*Si leo y empiezo a dudar, entonces no comparto*" (E2). Sin embargo, en general, afirmaron que no solían compartir noticias ("*Mis amigos y yo no compartimos muchas noticias*", E1). A la hora de contrastar la información, la mayoría de los más jóvenes refirieron utilizar YouTube como fuente de contraste, mientras que, en los grupos de mayor edad, Google apareció como la plataforma más utilizada para buscar información específica.

Cuando se preguntó a los participantes en la encuesta sobre cómo valoraban el grado de importancia de distintos aspectos a la hora de compartir noticias o contenidos online (Figura II), el elemento más señalado fue creer que la información procede de un autor o medio conocido o que se recibe de una persona de confianza. También se señaló contrastar la información antes de compartirla. En menor medida, aunque también relevante, apareció que la información sea divertida y que esté en formato vídeo o imagen. Lo que menos valoraron los encuestados fue el formato texto y que la información fuera potencialmente interesante para sus seguidores.

Temas y alcance de la difusión de noticias falsas entre los adolescentes

Mientras que en los grupos de Estonia y Eslovaquia los participantes mencionaron los tabloides como fuente de noticias falsas, en el resto de

FIGURA II. Grado de preocupación al compartir noticias (N=682)



Fuente: Elaboración propia.

Recibido: 31/03/2023 Aceptado: 22/12/2023

los grupos sólo se señalaron las redes sociales. Los temas más comunes apuntados en torno a las noticias falsas fueron, en primer lugar, la desinformación relacionada con la salud. El primer contacto con las fake news para algunos de los participantes se dio durante los primeros meses del COVID-19: "*Mi madre me dijo que había una vacuna para el COVID-19, y después me dijo que era un bulo. Me impresionó bastante porque no las había visto* [las noticias falsas sobre ningún tema] *antes"* (SP1); "*Como las noticias sobre que el hospital estaba vacío, etc*." (E2). Otros temas de salud señalados incluyeron supuestos riesgos para la salud, como la radiación cósmica que emanaría de nuestros teléfonos mientras dormimos.

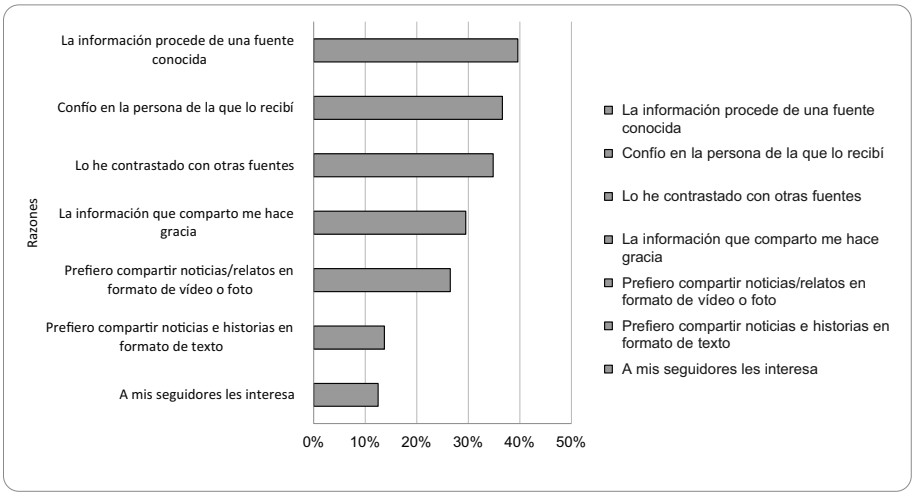
Un segundo tema destacado fueron los rumores sobre famosos (por ejemplo, romances o muertes falsas) y, en tercer lugar, diferentes noticias alarmantes que generan pánico social (por ejemplo, "*el día de la violación"* SP2, P1, "*secuestros*" (SP2)).

En cuanto a la encuesta, en la Figura III se puede observar que el 27,9% de la muestra (n = 190) reconoció haber compartido noticias o historias y enterarse después de que eran falsas (difusión no intencionada), y una cantidad ligeramente mayor (33,9%, n = 231) conoce a alguien de su edad que lo ha hecho (testigo de difusión no intencionada entre iguales). Aunque en menor medida, el 8,9% de la muestra (n = 61) reconoció haber compartido noticias falsas a sabiendas de que lo eran (difusión intencionada), y el 16,3% (n = 126) conocía a alguien que también lo hizo a propósito (testigo de difusión intencionada entre iguales).

En la encuesta se observaron diferencias significativas por edades en cuanto a la difusión de noticias falsas. Alrededor del doble de los adolescentes de más edad (15-17) en comparación con los más jóvenes (13-14) reconocieron haber compartido noticias falsas sin saber que lo eran (χ^2 = 8,824, p= .003, CC=.113), así como hacerlo conociendo su falsedad antes de compartirlas (χ^2 = 4,970, p = .021, CC=.085). Además, los estudiantes de la muestra de la encuesta declararon conocer a más personas de su edad que compartían noticias tanto sin saberlo (33,9%, n = 231) como a sabiendas (18,5%, n = 126) que personas mayores que hicieran lo mismo tanto sin saberlo (25,7%, n = 175) como a sabiendas (14,4%, n = 98).

Recibido: 31/03/2023 Aceptado: 22/12/2023

FIGURA III. Porcentaje de jóvenes que han compartido o presenciado noticias falsas en Internet



Fuente: Elaboración propia.

Adolescentes que difunden noticias falsas: sin saberlo y de forma intencional

Como se ha descrito anteriormente, las principales noticias que los participantes en los grupos de discusión recuerdan haber recibido o incluso compartido estaban relacionadas con sus intereses (cotilleos, conciertos, etc.) o tenían algún contenido emocional (secuestros y violaciones). De hecho, cuando se les pregunta por las razones de la difusión de noticias falsas, sugieren que es "*para llamar la atención de la gente y apelar a sus intereses. Por eso se difunden más rápido*" (P1). Además, los participantes describieron el papel que desempeña la diversión: un artículo que suena divertido se comparte mucho más fácilmente: "*Suelo compartir* [el contenido] *sí me parece divertido. A veces sólo leo el titular y, si me hace gracia, lo comparto*" (E2). En particular, incluso cuando el receptor sabe que el contenido divertido no es cierto, puede tener un impacto en la persona, como explicó otro participante: "*A veces leo las noticias en Internet porque son divertidas y, aunque no parezcan ciertas, me las creo un poco*" (E2).

Recibido: 31/03/2023 Aceptado: 22/12/2023

Según los participantes, las motivaciones de terceros para compartir noticias falsas serían ganar popularidad y fama ("*Las personas que buscan activamente conseguir el mayor número posible de likes, menciones, etc. tienden a compartir noticias falsas*" (G1)); sólo por diversión y entretenimiento o porque son "*personas incultas*" (E2). Por último, los participantes también afirmaron que muchos difusores creen que las noticias son ciertas, y su motivación es informar o ayudar a los demás.

En la encuesta, encontramos varias características, tanto en los difusores intencionados como en los no intencionados de noticias falsas, que eran significativamente diferentes a las del resto de participantes de la muestra general. En primer lugar, encontramos una tendencia significativa entre las diferentes motivaciones que los participantes expresan a la hora de compartir contenidos. Por un lado, los motivos que fueron significativamente más destacados entre quienes compartieron noticias falsas sin saber que eran falsas fueron que la información procediera de un periódico/autor conocido y que la información hubiera sido contrastada previamente por ellos mismos. Por otro lado, uno de los motivos para compartir contenidos de los participantes que compartieron noticias falsas sabiendo que eran falsas fue que la información fuera divertida. Además, entre estos participantes encontramos una correlación inversa respecto a que la información procediera de un periódico/autor conocido y que hubiera sido contrastada previamente (ver Tabla III). En segundo lugar, otros factores que arrojaron diferencias significativas fueron la edad, la tecnología utilizada y si existía algún antecedente migratorio (en el caso de los difusores no intencionados). También se mostró significativo haber recibido menos información y menos supervisión en casa por parte de sus progenitores (para los difusores intencionados) y reconocerse como agresor de discurso de odio online (sexismo, racismo y LGTBIfobia), como se muestra en la Tabla III.

Noticias falsas y discurso de odio

En los grupos de discusión se señaló una relación entre el ciberodio y las noticias falsas (por ejemplo, "*el ciberodio está relacionado con la difusión de información falsa*" (SL1)). Mostrar desinterés fue la reacción más frecuente ante el ciberodio ("*lo ignoro*" (P1 y SL1); "*no me hace nada*" (SL1)), incluso si sus características coincidían con las de los

TABLA III. Resumen de las características de los difusores y cuestiones que consideran importantes a la hora de compartir contenidos

	Compartir sin intención[1]	**Compartir a propósito**
Edad Mayor (15-17) vs. Menor (12-14)	**15-17 y-o (χ^2 = 8.7, p = .003, CC = .112)**	**15-17 y-o (χ^2 = 4.63, p =.031, CC = .082)**
Género Chicas vs. Chicos	(χ^2 = 4.509, p = .212, CC = .081)	(χ^2 = 7.223, p = .065, CC = .102)
Tipo de centro educativo Privado vs. Público	(χ^2 = 1.08, p = .582, CC = .040)	(χ^2 = .607, p = .738., CC = .030)
Origen migrante Sí vs. No	**Sí (χ^2 = 10.47, p = .015, CC= .123)**	(χ^2 = 1.17, p = .759, CC = .041)
Uso de Smartphone Alto vs. Bajo uso	**Uso alto (χ^2 = 14.647, p = .005, CC = .145)**	(χ^2 = 5.12, p = .271, CC = .087)
Uso de Tablet Alto vs. Bajo uso	**Uso alto (χ^2 = 14.917, p = .005, CC =.146)**	(χ^2 = 1.41, p = .843, CC = .045)
Uso de ordenador Alto vs. Bajo uso	(χ^2 = 8.85, p = .065, CC = .113)	(χ^2 = 3.91, p = .418, CC = .076)
Horas en línea durante la semana	(χ^2 = 6.98, p = .137, CC = .101)	(χ^2 = 2.58, p = .630, CC = .062)
Horas en línea durante el fin de semana	(χ^2 = 7.62, p = .107, CC = .105)	(χ^2 = 8.9, p = .064, CC = .114)
Información sobre los riesgos en casa Alta vs. Baja	(χ^2 = 3.78, p = .436, CC = .074)	**Baja (χ^2 = 13.27, p = .010, CC = .138)**
Supervisión parental Alta vs. Baja	(χ^2 = 5.22, p = .266, CC = .087)	**Baja (χ^2 = 10.976, p = .027, CC = .126)**
Comportamientos sexistas en Internet	(χ^2 = .359, p = .549, CC =.023)	**Alto riesgo (χ^2 = 37.92, p < .001, CC = .229)**
Comportamiento LGTBIfóbico en línea	(χ^2 = .146, p = .703, CC = .015)	**Alto riesgo (χ^2 = 35.12, p < .001, CC = .222)**
Comportamiento racista en línea	(χ^2= .177, *p* = .674, *CC* = .016)	**Alto riesgo (χ^2 = 27.76, *p* < .001, *CC* = .198).**
Relevancia atribuida a algunos factores a la hora de compartir información/noticias		
La información es divertida	(χ^2 = 8.63, p = .125, CC = .112)	**Mayor importancia (χ^2 = 32.08, p < .001, CC = .212)**

(Continuada)

Recibido: 31/03/2023 Aceptado: 22/12/2023

TABLA III. Resumen de las características de los difusores y cuestiones que consideran importantes a la hora de compartir contenidos *(Continuada)*

	Compartir sin intención[1]	**Compartir a propósito**
La información procede de un periódico/autor conocido	**Mayor importancia (χ^2 = 18.07, p = .003, CC = .161)**	**Menor importancia (χ^2 = 16.61, p = .005, CC = .154)**
La información ha sido contrastada previamente	**Mayor importancia (χ^2 = 12.055, p = .015, CC = .145)**	**Menor importancia (χ^2 = 13.12, p = .022, CC = .137)**

Fuente: Elaboración propia. Nota: los resultados significativos están marcados en negrita, el coeficiente de contingencia se expresa con "CC".

atacados: dos participantes negras señalaron sobre los contenidos racistas que "*depende del día, pero muchos días, paso... Porque no me apetece discutir*" (SP1). Aunque minoritarios, algunos participantes sí bloquearon o denunciaron los contenidos de ciberodio.

Los jóvenes también describieron cómo el ciberodio se disfraza a menudo de humor, como explicó un participante:

> *"En Tik Tok hay mucho humor con el tema del sexismo y todo eso, con la homofobia, el racismo y todo eso. Entonces, si alguien comenta algo como "esto es sexismo", todo el mundo le dice: es humor, no tenéis humor y eso. Nadie lo ve como sexismo sino como una broma" (SP2)*

En cuanto a la encuesta, como se ha visto en la sección anterior, el ciberodio (sexismo, racismo y LGTBIfobia) fue llevado a cabo con frecuencia por difusores intencionados de noticias falsas. También es significativo que, mientras que la preocupación general por las noticias falsas es bastante baja entre todos los participantes (tanto los adolescentes de los grupos de discusión europeos como los de la encuesta española, -donde el 58,21% estaban poco o nada preocupados por ello-), los tres grupos potenciales de victimización por ciberodio parecen ser más conscientes del fenómeno que otros participantes: las chicas (χ^2 = 29,34, p = .015, CC = .203), los migrantes (χ^2 = 27,05, p = .028, CC = .195) y los participantes LGTBI (χ^2 = 25,11, p = .048, CC = .188).

Discusión

Aunque investigaciones anteriores han demostrado cómo los adolescentes tienden a desconfiar de los medios de comunicación tradicionales (Spilker et al., 2020), por el contrario, la mayoría de los participantes de nuestra muestra coincidió en considerarlos como una fuente de información más fiable que las redes sociales. Aunque las prácticas y los discursos resultaron bastante similares en los distintos países, sólo en dos de ellos (Eslovaquia y Estonia) se señaló a un tipo específico de medios tradicionales -los tabloides- como fuente de noticias falsas. Así pues, la fiabilidad de los medios de comunicación podría depender del tipo de medios que son populares en los distintos países. No obstante, no se encontraron otras diferencias notables entre regiones.

A pesar de su mayor fiabilidad, los adolescentes señalaron que casi no consumen medios de comunicación tradicionales y que se informan principalmente a través de las redes sociales, como se ha constatado en otros estudios (por ejemplo, Pérez-Escoda et al., 2021). En consecuencia, más que discutir si los jóvenes desconfían de los medios tradicionales, las investigaciones futuras pueden centrarse en encontrar las razones de su escaso uso. Tal vez el lenguaje técnico disuade a los jóvenes de disfrutar de los medios tradicionales y podrían proponerse secciones con un lenguaje más amigable y accesible para atraer a los jóvenes lectores (de hecho, algunas revistas científicas ya se acercan al público más joven, véase Frontiers for Young Minds como ejemplo). Además, los participantes en nuestra muestra describieron las redes sociales como lugares donde se encuentra información de última hora y controvertida, con opiniones diferentes que suelen aprovechar para formar su propia visión de la actualidad (Marchi, 2012) y que congregan "a todo el mundo allí". Esta descripción parece estar relacionada con cierta urgencia en los adolescentes por buscar a los demás, interactuar con ellos y crear una identidad propia, lo que encaja con las características de la generación Z (Mendiguren et al., 2020). En este sentido, los medios de comunicación tradicionales no conectarían tanto con los intereses de los jóvenes como lo hacen las redes sociales.

También hemos comprobado que, si bien para ser denominado "fake news" el contenido tiene que ser creado deliberadamente (Shu et al., 2017), el difusor no necesariamente comparte con el creador este objetivo engañoso a propósito, ya que la mayoría de los difusores no sabrían

Recibido: 31/03/2023 Aceptado: 22/12/2023

que están compartiendo contenidos falsos. En la muestra de la encuesta, mientras que casi un tercio de los participantes había compartido noticias falsas sin saberlo -un porcentaje menor que en otras investigaciones y sin diferencias de género (Chen et al., 2015)-, menos del 10% lo hizo a propósito.

Las características que los grupos de discusión señalaron para que las noticias falsas tuvieran mayor probabilidad de ser creídas y compartidas, coincidían con las características que la investigación ha destacado para que una noticia resulte atractiva, como contenido emocional, relevancia para el receptor y emisores bien informados (Duffy et al., 2020), además de ciertos tipos de contenido como historias divertidas e información surrealista y sorprendente que, según investigaciones anteriores, también importan a otros adolescentes (Baptista, 2020; Herrero-Diz et al., 2020). En definitiva, los adolescentes parecen preferir contenidos emocionales, llamativos y relevantes para sus intereses (Herrero-Diz et al., 2020), por lo que esto podría explicar que nuestra muestra no considerase fake news relacionadas con la política -como generalmente se encuentra (Goyanes & Lavin, 2018)- sino más relacionadas con temas de salud, vida de famosos (Gómez Calderón et al., 2020) y pánicos morales, temas que les interesan más. Este hallazgo es especialmente relevante a efectos de prevención, ya que, previsiblemente, si se utiliza este tipo de contenidos para llamar la atención y enseñar a los jóvenes a detectar la información falsa, la intervención será más efectiva. El éxito en Internet de las noticias falsas sobre salud, especialmente las relativas a la COVID-19, demuestra que responder a las dudas de la gente funciona eficazmente, es decir, que cuando la gente está preocupada por algún tema, buscará respuestas en Internet, terminando a veces en una espiral de información sin contrastar, impulsada por los algoritmos de las redes sociales.

Sin embargo, y de forma llamativa, aunque investigaciones anteriores han sugerido que el intercambio de noticias sería una parte importante de la forma en que los jóvenes expresan su identidad y mantienen la conexión social (Marchi, 2012), en nuestra muestra la mayoría de los participantes afirmaron no compartir tantas noticias. Además, la mayoría declaró leer principalmente sólo las noticias que aparecían en sus redes sociales publicadas por amigos. Esta podría ser la razón por la que, en términos generales, no les preocupan especialmente las noticias falsas. Su difusión a través de Internet se entiende como poco preocupante tanto en los grupos de discusión como en la encuesta, muy por detrás

de otros riesgos de Internet como el ciberacoso y el online grooming. En este sentido, los programas educativos deberían incidir en los peligros y los efectos de las noticias falsas, al no ser consecuencia el problema -o no sólo- de una cuestión de falta de destreza a la hora de detectarlas sino, sobre todo, de subestimar su efecto.

En cuanto a la forma en que los adolescentes comparten las noticias según la encuesta, se prefirió el formato de vídeo e imagen al de texto (como en Literat et al., 2020), a pesar de la aparición y popularidad de nuevos softwares que permiten producir deep fakes (por ejemplo, imitando discursos reales dentro de fuentes de vídeo no originales; creando imágenes con apariencia real pero que nunca sucedieron en el mundo real) o del uso generalizado de programas de edición de imágenes, normalmente ya incorporados a los smartphones que los adolescentes utilizan habitualmente. Sin embargo, aunque la veracidad de la información y la autoridad de la fuente no resultaron relevantes para compartir información en investigaciones anteriores (Chen et al., 2015), en nuestra muestra estos factores fueron los considerados más importantes, lo que podría indicar cierto cambio positivo en la tendencia. No obstante, la confianza en el emisor también fue señalada como una variable relevante a la hora de decidir si compartir o no información, lo que puede llegar a ser arriesgado, especialmente si tenemos en cuenta que su principal canal de información son las redes sociales. Por un lado, en éstas abundan los sujetos de confianza. Por otro, la doble verificación puede verse comprometida al ser interpretada como un gesto de desconfianza. Parece que habría que educar a los adolescentes en el doble control e informarles de cómo la difusión de noticias falsas puede provocar que incluso personas de confianza se conviertan también en emisores de desinformación. En una línea igualmente arriesgada, los participantes también afirmaron confiar en el diseño y el refinamiento del contenido, como se ha destacado en investigaciones anteriores (McGrew et al., 2017). Dado que muchos sitios web de contenido dudoso son cada vez más sofisticados, habría que educar a los jóvenes sobre cómo cambiar de un sitio web a otro para contrastar los datos, en lugar de seguir un criterio estético.

Por último, una de las principales estrategias que los adolescentes reportaron realizar a la hora de decidir sobre la credibilidad de algún contenido fue revisar los comentarios de la publicación, como también se ha encontrado previamente en otros estudios (Colliander, 2019). Aunque

comprobar los comentarios puede considerarse una forma de pensamiento crítico, no está exento de riesgos, ya que puede encontrarse fácilmente en estrategias de marketing, como los comentarios patrocinados. Esta tendencia, junto con la preferencia por noticias con opiniones y temas controvertidos, encajaría en la descripción de Daum (2019) de las nuevas generaciones que utilizarían menos el pensamiento crítico, persiguiendo un pensamiento más pasivo de dentro/fuera del grupo. Esto quiere decir que el sujeto asume su posición sobre determinados temas tras leer lo que piensan los demás. Por ello, ser escéptico ante los comentarios debería incluirse en los programas preventivos de desinformación para adolescentes.

Respecto a los factores de riesgo relacionados con la difusión no intencionada de noticias falsas entre los adolescentes, se encontró tener origen migrante independientemente del país de origen (es decir, no sólo los procedentes de países de renta baja). Esto podría estar relacionado con la dificultad para comprender con precisión los contenidos que encuentran en Internet debido a sus limitaciones lingüísticas. Además, los difusores no intencionados utilizaban más los teléfonos móviles y las tabletas para navegar por Internet. Esta preferencia podría estar relacionada con el hecho de que los dispositivos de mano favorecen la impulsividad, característica asociada a la difusión de contenidos falsos (Herrero-Diz et al., 2020), ya que dificultan comprobar la veracidad del contenido frente a hacerlo a través de un ordenador. A diferencia de lo encontrado en otros estudios (Goyanes, & Lavín, 2018), podemos sugerir indirectamente que los ingresos familiares no fue un factor relevante, ya que no hubo diferencias entre los distintos tipos de centros educativos.

En cuanto a la importancia otorgada a distintos aspectos a la hora de compartir información, mientras que los divulgadores intencionados valoraron la fuente y el hecho de que la información estuviera contrastada en menor medida que el conjunto de la muestra, los divulgadores no intencionados, por el contrario, la valoraron por encima del resto. Esto podría deberse precisamente al hecho de haber sido víctimas involuntarias de las fake news.

También se encontró una relación entre difundir noticias falsas intencionalmente y convertirse en agresor de discurso de odio en Internet. Esto podría ilustrar cómo el uso ideológico de las noticias falsas está teniendo éxito entre los jóvenes. Además, no sólo la etnia y la nacionalidad fueron objetivos del ciberodio, sino también la orientación

sexual y el género (Castaño-Pulgarín et al., 2021). La ideología del discurso de odio podría producir un sesgo cognitivo en las personas que siguen a personas que difunden este tipo de contenidos, de forma que al recibir información, podrían estar más inclinados a creer información que confirme sus creencias previas (Weidner et al., 2020) y, en consecuencia, llevarles a no evaluar el contenido de forma objetiva. Los usuarios tienden a estar rodeados de personas con ideas afines, lo que significa que podrían estar expuestos sólo o principalmente a una parte de toda la historia que puede ser parcialmente cierta o totalmente falsa. Además, el algoritmo amplifica este fenómeno creando burbujas de filtros, de forma que cuando alguien ve una publicación o le gusta, registra que la persona está interesada en esa información y se le mostrarán más contenidos relacionados con ella en el futuro. La motivación ideológica, además, podría provocar el hecho de compartir la noticia incluso sabiendo que la información de la misma no es veraz.

Por lo tanto, las medidas preventivas en torno a las noticias falsas no sólo deben abordar la alfabetización mediática, sino también la educación contra el discurso del odio y la discriminación. Igualmente, se debería incluir a las familias en la prevención, ya que recibir poca información y poca supervisión en casa son factores de riesgo. Asimismo, hay cierta trivialización de las noticias falsas y del discurso del odio, ya que el humor parece desempeñar un papel tanto en la justificación para compartir noticias falsas como en cierta forma de enmascaramiento del discurso de odio. Esto apunta a la necesidad de dotar a los jóvenes de herramientas para el empoderamiento, bajo la forma de contra-discurso (Blaya, 2019).

Nuestros resultados y conclusiones no están exentos de limitaciones. Para profundizar en el fenómeno, utilizamos dos metodologías diferentes, pero debido a la dimensión del proyecto, que involucraba a varios países, las muestras no estaban equilibradas entre países, y la triangulación con la encuesta sólo se realizó en un país (España). Sin embargo, considerando que no encontramos diferencias relevantes entre países y que una comparación interregional no era nuestro objetivo, entendemos además que las fake news son un fenómeno global entre adolescentes. Aunque se trata de un estudio exploratorio, nuestros resultados, tanto en los grupos de discusión como en la encuesta son bastante consistentes, para ser tenidos en cuenta como información preliminar que puede permitir una investigación más profunda. Aportan además elementos útiles para ser tenidos en cuenta en las intervenciones educativas y los proyectos de prevención.

Recibido: 31/03/2023 Aceptado: 22/12/2023

Conclusiones

En resumen, pese a las limitaciones del estudio, los resultados obtenidos apuntan una serie de cuestiones relevantes a la hora de diseñar medidas o programas de prevención. Primero, habría que hacer hincapié en las consecuencias de distribuir y consumir ese contenido, ya que no existe un desconocimiento sino más bien una despreocupación. Segundo, para que las intervenciones tengan éxito deberían incluir como ejemplos contenidos emocionales, llamativos y relacionados con los intereses de los jóvenes. Tercero, habría que llamar la atención respecto a tres criterios poco fiables de veracidad que suelen considerar válidos: la afinidad o la confianza hacia la persona que comparte el contenido (se podría poner como ejemplo que es un amigo íntimo el que, sin saberlo, comparte una noticia falsa); el grado de sofisticación del diseño de la fuente (de nuevo, se pueden utilizar ejemplos con un buen diseño, sin erratas y apariencia científica); y la contrastación a través de los comentarios. Finalmente, junto a la alfabetización mediática se debería abordar la difusión del discurso de odio y el peligro de la banalización de los mensajes discriminatorios a través del humor.

Financiación

Este estudio se deriva del proyecto RAYUELA (Ref: H2020-SU-SEC-2019-882828), Empowering and educating young people for the internet by playing, que ha recibido financiación del programa de investigación e innovación Horizonte 2020 de la Unión Europea en virtud del acuerdo de subvención nº 882828. El contenido del estudio refleja únicamente la opinión de los autores. La Comisión Europea o su Agencia Ejecutiva de Investigación no son responsables del uso que pueda hacerse de la información que contiene.

Referencias bibliográficas

Aguaded, J.I. (2012). United Nations aiming at Media Literacy Education. *Comunicar, 38*, 7-8. http://dx.doi.org/10.3916/C38-2012-01-01

Al-Nasser, N. A. (2017). Media and information literacy as a means of preventing violent extremism. *UN Chronicle*, *54*(3), 30-31. Recuperado de: https://www.un.org/en/chronicle/article/media-and-information-literacy-means-preventing-violent-extremism

Anderau, G. (2021). Defining fake news. *KRITERION–Journal of Philosophy, 35*(3), 197-215. https://doi.org/10.1515/krt-2021-0019

Aparici, R. (2011). *Educomunicación: más allá del 2.0*. Gedisa

Apuke, O. D., & Omar, B. (2020). Modelling the antecedent factors that affect online fake news sharing on COVID-19: the moderating role of fake news knowledge. *Health Education Research*, *35*(5), 490-503. https://doi.org/10.1093/her/cyaa030

Baptista, J. P., & Gradim, A. (2020). Understanding fake news consumption: A review. *Social Sciences*, *9*(10), 185. https://doi.org/10.3390/socsci9100185

Blaya, C. (2019). Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior*, *45*, 163-172. https://doi.org/10.1016/j.avb.2018.05.006

Castaño-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. (2021). Internet, social media and online hate speech. Systematic review. *Aggression and Violent Behavior*, *58*, 1-7. https://doi.org/10.1016/j.avb.2021.101608

Cerase, A., & Santoro, C. (2018). 15. From racial hoaxes to media hypes. In P. Vasterman (Ed.), *From Media Hype to Twitter Storm: News Explosions and Their Impact on Issues, Crises and Public Opinion* (pp. 333-354). University Press. https://doi.org/10.1515/9789048532100-018

Chen, X., Sin, S. C. J., Theng, Y. L., & Lee, C. S. (2015). Why students share misinformation on social media: Motivation, gender, and study-level differences. *The journal of academic librarianship*, *41*(5), 583-592. https://doi.org/10.1016/j.acalib.2015.07.003

Colliander, J. (2019). "This is fake news": Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media. *Computers in Human Behavior*, *97*, 202-215. https://doi.org/10.1016/j.chb.2019.03.032

Daum, M. (2019). *The problem with everything: My journey through the new culture wars*. Gallery Books.

Duffy, A., Tandoc, E., & Ling, R. (2020). Too good to be true, too good not to share: the social utility of fake news. *Information, Communication & Society*, *23*(13), 1965-1979. https://doi.org/10.1080/1369118X.2019.1623904

Recibido: 31/03/2023 Aceptado: 22/12/2023

Fernández-García, N. (2017). Fake news: una oportunidad para la alfabetización mediática. *Nueva Sociedad, 269*, 66-77. Retrieved from https://bit.ly/2MpoOPV

Gil Flores, J., García Jiménez, E., & Rodríguez Gómez, G. (1994). El análisis de los datos obtenidos en la investigación mediante grupos de discusión. *Enseñanza and Teaching, 12*(17), 183-199. Retrieved from: http://hdl.handle.net/10366/69459

Gómez Calderón, B., Córdoba-Cabús, A., & Méndez-Nieto, A. (2020). Jóvenes y fake news. Un análisis sociodemográfico aplicado al caso andaluz. *IC: Revista Científica de Información y Comunicación, 17,* 481-504. https://doi.org/10.12795 / IC.2020.i01.21

Goyanes, M., & Lavin, A. (2018). The Sociology of Fake News–Factors affecting the probability of sharing political fake news online. *Media@ LSE Working Paper Series*, *20*.

Gozálvez, V., y Contreras-Pulido, P. (2014). Empoderar a la ciudadanía mediática desde la educomunicación. Comunicar. *Revista Científica de Educomunicación, 42*(XXI), 129-136. https://doi.org/10.3916/C42-2014-12

Hawdon, J., Oksanen, A., & Räsänen, P. (2015). Online extremism and online hate. *Exposure among Adolescents and Young Adults in Four Nations. NORDICOM*, *3*, 29-37. Retrieved from: https://www.diva-portal.org/smash/get/diva2:939050/FULLTEXT01.pdf

Herrero-Diz, P., Conde-Jiménez, J., & Reyes de Cózar, S. (2020). Teens' motivations to spread fake news on WhatsApp. *Social Media+ Society*, *6*(3), 1-14. http://doi.org/10.1177/2056305120942879

Herrero-Diz, P., Conde-Jiménez, J., & Reyes-de-Cózar, S. (2021). Spanish adolescents and fake news: level of awareness and credibility of information (Los adolescentes españoles frente a las fake news: nivel de conciencia y credibilidad de la información). *Culture and Education*, *33*(1), 1-27. https://doi.org/10.1080/11356405.2020.1859739

Literat, I., Chang, Y. K., & Hsu, S. Y. (2020). Gamifying fake news: Engaging youth in the participatory design of news literacy games. *Convergence*, *26*(3), 503-516. https://doi.org/10.1177/1354856520925732

Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity". *Journal of Communication Inquiry*, *36*(3), 246-262. https://doi.org/10.1177/0196859912458700

Marwick, A. E. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review*, *2*(2), 474-512. Retrieved from: https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-512.pdf

Mascheroni, G., & Ólafsson, K. (2014). *Net children go mobile: Risks and opportunities*. Retrieved from: https://eprints.lse.ac.uk/55798/1/Net_Children_Go_Mobile_Risks_and_Opportunities_Full_Findings_Report.pdf

McGrew, S., Ortega, T., Breakstone, J., & Wineburg, S. (2017). The Challenge That's Bigger than Fake News: Civic Reasoning in a Social Media Environment. *American Educator*, *41*(3), 4-9.

Mendiguren, T., Pérez Dasilva, J., & Meso Ayerdi, K. (2020). Actitud ante las Fake News: Estudio del caso de los estudiantes de la Universidad del País Vasco. *Revista de Comunicación*, *19*(1), 171-184. http://doi.org/10.26441/rc19.1-2020-a10

Muigai, J. W. (2017). Understanding fake news. *Washington Post*. http://doi.org/10.29322/IJSRP.9.01.2019.p8505

Páramo, M. (2011). Factores de Riesgo y Factores de Protección en la Adolescencia: Análisis de Contenido a través de Grupos de Discusión. *Terapia psicológica*, *29*(1), 85-95. http://doi.org/10.4067/S0718-48082011000100009

Pérez-Escoda, A., Pedrero-Esteban, L. M., Rubio-Romero, J., & Jiménez-Narros, C. (2021). Fake news reaching young people on social networks: Distrust challenging media literacy. *Publications*, *9*(2), 24. https://doi.org/10.3390/publications9020024

Schäfer, C., & Schadauer, A. (2018). Online fake news, hateful posts against refugees, and a surge in xenophobia and hate crimes in Austria. In *Refugee News, Refugee Politics* (pp. 109-116). Routledge.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, *19*(1), 22-36. https://doi.org/10.1145/3137597.3137600

Stabile, B., Grant, A., Purohit, H., & Harris, K. (2019). Sex, lies, and stereotypes: Gendered implications of fake news for women in politics. *Public Integrity*, *21*(5), 491-502. https://doi.org/10.1080/10999922.2019.1626695

Spilker, H. S., Ask, K., & Hansen, M. (2020). The new practices and infrastructures of participation: How the popularity of Twitch. tv challenges old and new ideas about television viewing. *Information, Communication & Society*, *23*(4), 605-620. https://doi.org/10.1080/1369118X.2018.1529193

Recibido: 31/03/2023 Aceptado: 22/12/2023

Tompson, T. (2017). *Paying for News: Why People Subscribe and What It Says About the Future of Journalism*, The Associated Press-NORC Center for Public Affairs Research. Retrieved from https://policycommons.net/artifacts/2148307/paying-for-news/290380/

UNESCO (2009). *Informe sobre las actividades del Consejo Intergubernamental del Programa Internacional para el Desarrollo de la Comunicación (PIDC) (2008-2009)*. Recuperado de: https://unesdoc.unesco.org/ark:/48223/pf0000183419_spa

Verstraete, M., Bambauer, D. E., & Bambauer, J. R. (2018). Identifying and Countering Fake News. *SSRN Electronic Journal*, 17.

Vziatysheva V. (2020). How fake news spreads online? *International Journal of Media and Information Literacy*, *5*(2), 217-226. http://doi.org/10.13187/ijmil.2020.2.217

Weidner, K., Beuk, F., & Bal, A. (2020). Fake news and the willingness to share: a schemer schema and confirmatory bias perspective. *Journal of Product & Brand Management*, *29*(2), 180-187. https://doi.org/10.1108/JPBM-12-2018-2155

Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2020). Hate in the machine: Anti-Black and anti-Muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*, *60*(1), 93-117. https://doi.org/10.1093/bjc/azz049

Información de contacto: María Reneses. Universidad Pontificia Comillas C. Universidad Comillas, 28049 Alcobendas, Madrid. E-mail: mariareneses@gmail.com

Hábitos de vida saludables en educación primaria: revisión sistemática

Healthy lifestyle habits in primary education: a systematic review

https://doi.org/10.4438/1988-592X-RE-2024-406-646

Olatz Arce-Larrory
https://orcid.org/0000-0003-1412-332X
Universidad de Deusto

Erlantz Velasco
https://orcid.org/0000-0002-4554-9852
Universidad de Deusto

Iker Sáez
https://orcid.org/0000-0002-6094-2812
Universidad de Deusto

Resumen

El entorno escolar tiene el potencial de influir tanto en la salud como en el bienestar del alumnado. En este sentido, han sido numerosos los estudios que analizan y evalúan intervenciones educativas para el fomento de los hábitos de vida saludables en Educación Primaria. Aun así, muchos, se han desarrollado fuera de las áreas curriculares, sin tener en cuenta el valor de estas en el fomento de dichos hábitos. Por tanto, el presente artículo tiene por objetivo analizar intervenciones realizadas para mejorar los hábitos de vida saludables en el alumnado de Educación Primaria en el espacio curricular entre los años 2015-2022. Para ello, se realizó una revisión sistemática siguiendo la declaración PRISMA. El proceso de selección de artículos arrojó un total de 24 artículos a analizar, en los cuales se realizó un análisis descriptivo y un análisis de contenidos. Las siguientes variables fueron analizadas: población, asignatura, agente que dirige

Recibido: 27/04/2023 Aceptado: 22/12/2023

la intervención, estructura de la intervención, tipo de actividades, formación, implicación de la familia, tareas adicionales, y recursos utilizados. Los resultados generales permiten confirmar la inexistencia de un marco declarativo sobre los aspectos o dimensiones que abarcan los hábitos de vida saludables en las intervenciones educativas de Educación Primaria. No obstante, del análisis se concluye que se observan elementos propios de aspectos relacionados con la actividad física, la nutrición, el bienestar emocional y la higiene.

Palabras clave: hábitos de vida saludables, educación primaria, currículum, programa educativo, intervención.

Abstract

The school environment has the potential to influence both the health and well-being of pupils. In this sense, many studies have analysed and evaluated educational interventions for the promotion of healthy lifestyle habits in primary education. However, many were developed outside the curricular areas, without considering their value in the promotion of these habits. Therefore, this article aims to analyse interventions carried out to improve healthy lifestyle habits in primary school pupils within the curricular space between 2015 and 2022. To this end, a systematic review was performed following the PRISMA statement. The article selection process yielded a total of 24 articles for analysis, of which descriptive and content analyses were conducted. The following variables were analysed: population, subject, agent directing the intervention, intervention structure, type of activities, training, family involvement, additional tasks, and resources used. The general results confirm the lack of a declarative framework regarding aspects or dimensions encompassing healthy lifestyle habits in educational interventions of primary education. Nevertheless, the analysis concludes that there are elements specific to aspects related to physical activity, nutrition, emotional well-being, and hygiene.

Keywords: healthy lifestyle habits, primary education, curriculum, educational program, intervention.

Introducción

La Comisión OMS-UNICEF-The Lancet señala la existencia de un problema global de salud y bienestar infantil, y aboga por situar el bienestar infantil en el centro de las políticas Objetivos de Desarrollo Sostenible (ODS) (Clark et al., 2020). Sobre ese problema global, las

Recibido: 27/04/2023 Aceptado: 22/12/2023

enfermedades no transmisibles o enfermedades de estilos de vida (Akseer et al., 2020, OMS, 2022) se sitúan como agentes fundamentales. Estas enfermedades, se originan en la infancia y ya desde esa etapa evolutiva vienen determinadas por factores como la inactividad física, una dieta poco saludable o los entornos insalubres o inseguros (OMS, 2022). La relación entre la influencia de las experiencias de la infancia y el estado de salud en el futuro está documentada (Hahnraths et al., 2021; Whitaker et al., 1997). Es por ello, que resulta fundamental promover durante los primeros años de vida lo que la literatura define como hábitos de vida saludables (HVS) para, así, desarrollarlos y adherirlos durante la vida adulta (Gámez-Calvo et al., 2022; Hahnraths et al., 2021; Invernizzi et al., 2019; Langford et al., 2014; Moscoso-Loaiza y Díaz-Heredia, 2018).

Los HVS son un modo de vida que reducen el riesgo de padecer enfermedades o morir prematuramente (OMS, 1999). La literatura no recoge de manera unificada cuales son aquellos hábitos que componen dicho modo de vida, aunque parece ser, que la actividad física (AF) y la nutrición son dos de los elementos indispensables (de Villier et al., 2016; Jimenez-Parra et al., 2022; López et al., 2021; OMS, 1999), evidenciando la necesidad de abordar este tema. Estos hábitos se conforman desde la infancia, requieren de aprendizaje y repetición, e incluso llegan a volverse inconscientes (Moscoso-Loaiza y Díaz-Heredia, 2018; Oberto et al., 2020). El hecho de producirse en la fase evolutiva infantil, hace que tanto la familia como la educación se conviertan en ejes determinantes de su perpetuación (Gámez-Calvo et al., 2022; Hung et al., 2015; OMS, 2021).

Desde 1985, la OMS, UNESCO y UNICEF comenzaron a estructurar el enfoque de las Escuelas Promotoras de la Salud (*Health-Promoting School)* (OMS, 2021), evidenciando la importancia que tiene el entorno escolar a la hora de influir en la salud y el bienestar del alumnado (Bonell et al., 2013; Langford et al., 2014). Los niños y las niñas pasan gran parte del día en la escuela y, tal y como revelan diversas investigaciones este periodo lectivo en el aula puede convertirse en tiempo muy útil para educar y adquirir los HVS (Hahnraths et al., 2021; Pulido-Gil et al., 2022), por ejemplo, a través de la promoción de programas escolares (Jones et al., 2020; Markov et al., 2020). En este sentido, los programas que promuevan los HVS deben ser de abordaje holístico, en el que se integren todos los agentes educativos: equipo directivo, profesorado y

familia (Gámez-Calvo et al., 2022; Harris y Cale, 2018; y Moscoso-Loaiza y Díaz-Heredia, 2018).

Por todo ello, se presenta una revisión sistemática para conocer las dimensiones que abarcan los HVS en el currículum de Educación Primaria (EP), y, desde un punto de vista pedagógico, para dar cuenta que tipo de intervenciones se llevan a cabo desde el periodo de adopción de los ODS (2015) hasta la fecha de realización del estudio (2022).

Método

La revisión sistemática siguió la declaración PRISMA 2020. La pregunta de investigación, la estrategia de búsqueda y los criterios de inclusión y exclusión se elaboraron mediante la misma declaración PRISMA (Page et al., 2021).

Criterios de elegibilidad

Mediante la estrategia PICO (Tabla I) establecida por el *National Institute for Health and Care Excellence* (Schardt et al., 2007) se planteó la pregunta de investigación: ¿Qué intervenciones (I) se han implementado para mejorar los hábitos de vida saludables (O) del alumnado en EP (P)? De acuerdo a esto, se diseñaron los criterios de inclusión y exclusión, y se tuvieron en cuenta los artículos con diseños cualitativos, cuantitativos y mixtos. En la misma se excluye el término comparación (O) debido a que no se busca la comparación de los estudios propiamente.

Recibido: 27/04/2023 Aceptado: 22/12/2023

TABLA I. Criterios de inclusión y exclusión formulados con la estrategia PICO

	Población	**Intervención**	**Resultados**
Criterios de inclusión	Alumnado de EP	Intervenciones sobre HVS en EP en el ámbito curricular que detallan su procedimiento y evaluación	Informan sobre los hallazgos obtenidos en algún HVS
Criterios de exclusión	• Estudiantes que no se encuentran en la educación primaria. • Estudiantes con algún tipo de enfermedad (diabetes, obesidad, asma, cáncer, parálisis cerebral, discapacidad intelectual…) • Padres y madres, profesorado, trabajadores de la escuela • Otras profesiones, jugadoras, adultos, personas de la tercera edad, refugiados, personas en orfanatos, mujeres.	• Intervenciones extracurriculares, cocurriculares, formaciones, análisis del estado • No intervenciones • Intervenciones sin procedimiento y evaluación • Validación de escalas y herramientas • Revisiones sistemáticas o metaanálisis. • Congresos • Intervenciones no dirigidas al alumnado • Intervenciones dirigidas a la prevención de la obesidad o sobrepeso	• No informan sobre los resultados sobre los HVS.

Fuente: elaboración propia.

Fuentes de información

La búsqueda de los estudios se llevó a cabo el 27 de octubre de 2022 en las bases de datos de Scopus, PubMed, Web of Science y ERIC. El inglés fue el idioma predominante (87,5%) por detrás del castellano (12,5%). Se realizaron búsquedas de artículos publicados en revistas, desde el año de creación de las ODS (2015) hasta el año del diseño de la estrategia de búsqueda (2022). Se excluyeron los capítulos de libros, las revisiones sistemáticas o metaanálisis, informes y actas de congresos científicos.

Estrategia de búsqueda

Mediante la estrategia PICO se estableció la pregunta de investigación que ha guiado la revisión sistemática. La ecuación de búsqueda fue adecuada

a las condiciones de cada base de datos, utilizando los operadores booleanos AND y OR (Tabla II) (Methley et al., 2014).

TABLA II. Palabras clave formuladas con la estrategia PICO

	[1] **Población**	[2] **Intervención**	[3] **Resultados**
Palabras clave	"primary school*" OR "primary grad*" OR "primary education*" OR "elementary school" OR "elementary education"	intervention* OR program* OR activit*	"healthy lifestyle*" OR "healthy habit*" OR "well?being" OR wellness
Búsquedas	**En Scopus:** TITLE/ABS/KEY [1] AND TITLE/ABS/KEY [2] AND TITLE/ABS/KEY [3] **En PubMed:** TITLE/ABSTRACT [1] AND TITLE/ABSTRACT [2] AND TITLE/ABSTRACT [3] **En WoS:** TOPIC(TS) [1] AND TOPIC(TS) [2] AND TOPIC(TS) [3] **En ERIC:** TITLE/DESCRIPTOR [1] AND TITLE/DESCRIPTOR [2] AND TITLE/DESCRIPTOR [3]		
Límites	**En Scopus:** AND PUBYEAR > 2014 AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"re")) **En PubMed:** AND (2015:2022[pdat]) **En WoS:** refined by: Publication year (2015-2022) and Document types (Article or Review) **En ERIC:** pubyearmin:2015		

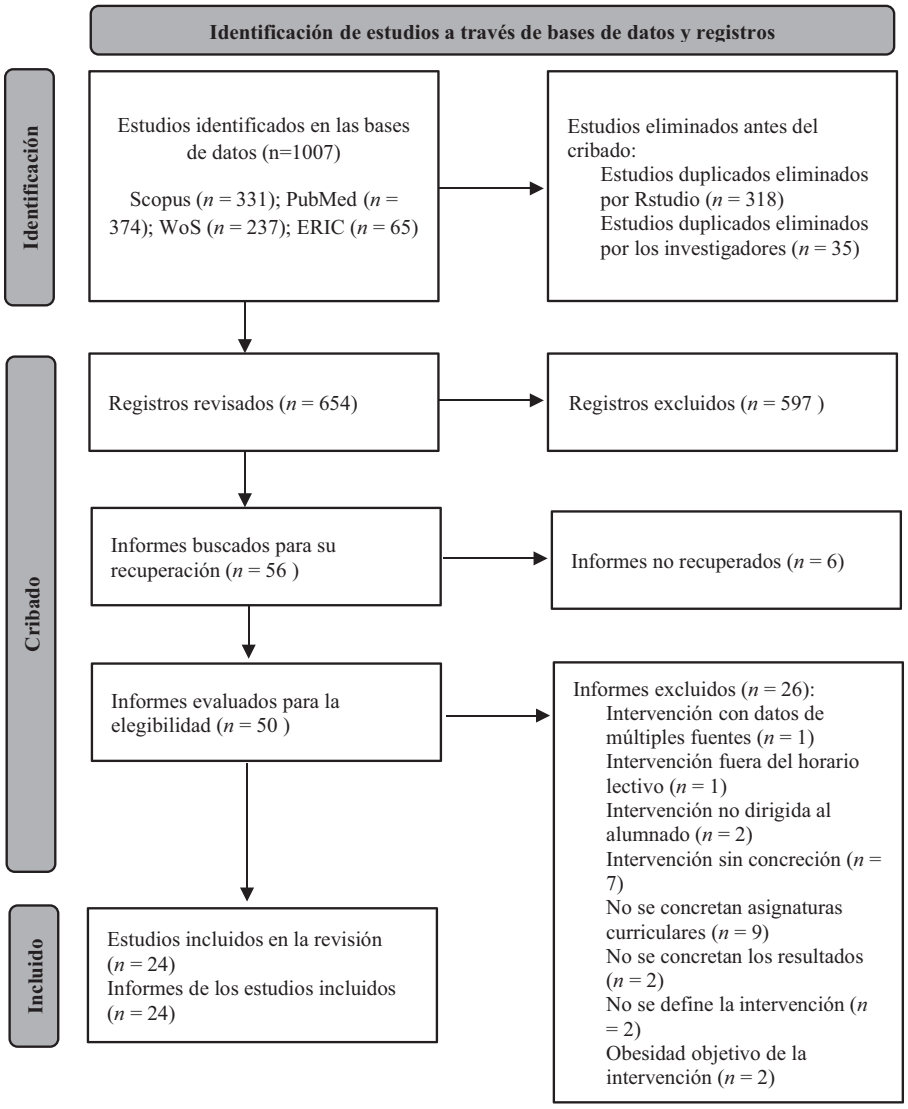
Fuente: elaboración propia.
Nota: [ ]: indica la numeración de las palabras clave utilizadas en la búsqueda. Estas se ven reflejadas en la búsqueda de cada base de datos.

Proceso de recogida de datos y selección

Los investigadores definieron los términos a utilizar en la estrategia PICO. La búsqueda fue revisada por los trabajadores del centro de recursos para el aprendizaje y la investigación (CRAI) de la Universidad de Deusto, para reafirmar la idoneidad de la búsqueda. Los estudios identificados fueron seleccionados por criterios de inclusión y exclusión (Tabla I). Los investigadores responsables llevaron a cabo el proceso de selección (Gráfico I). De manera independientemente, revisaron el cribado de los títulos, el resumen, y los estudios, en un documento Excel y, en caso de discrepancia, se resolvió mediante discusión hasta alcanzar un consenso de mayoría simple. Así, de los 1007 estudios identificados (654 excluyendo los duplicados) en las bases de datos Scopus, PubMed, WoS y ERIC, se seleccionaron 24 estudios a analizar.

Recibido: 27/04/2023 Aceptado: 22/12/2023

GRÁFICO I. Diagrama de flujo del procedimiento de selección de estudios



Fuente: Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ British Medical Journal, 372*(71), 1-9. http://doi.org/10.1136/bmj.n71

Recibido: 27/04/2023 Aceptado: 22/12/2023

Los documentos duplicados se eliminaron mediante Rstudio (R.4.2.0) (*n* = 318) y fueron revisados y excluidos por los investigadores del estudio de manera independiente (*n* = 35). Este proceso de selección siguió los criterios de inclusión y exclusión (Tabla I).

La investigadora principal analizó los estudios seleccionados y revisados por los otros dos investigadores. El análisis de contenido fue realizado mediante el programa ATLAS.ti (22.1.5.0) y la hoja de cálculo de Excel.

Elementos de datos

Las variables para las que se buscaron datos se detallan a continuación (Tabla III).

TABLA III. Las variables para las que se buscaron datos

Variables	**Descripción de las variables**
1. Tipo de artículo	Cuantitativo, cualitativo o mixto
2. El año	2015-2022
3. La revista	Nombre de la revista
4. El país	El país donde se implementó la intervención
5. El objetivo	Descripción del objetivo del estudio
6. La muestra	La población del estudio
7. La intervención del estudio: (1) Descripción general; (2) tipo de actividades; (3) requería formación o no; (4) requería la implicación de las familias; (5) incluía tareas adicionales	(1) Breve descripción de la intervención (2) Tipo de actividades llevadas a cabo (3) Si los agentes recibían formación o no (4) Si las familias formaban parte de las intervenciones (5) Si en las intervenciones se incluían tareas adicionales o no
8. La asignatura(s) donde se llevaba a cabo la intervención	Integración curricular (varias asignaturas curriculares) Asignatura curricular (una única asignatura curricular)
9. El agente	La persona implicada en la implementación de la intervención
10. Los recursos	Los recursos utilizados durante la intervención
11. Los resultados	Breve descripción de los principales resultados

Fuente: elaboración propia

Recibido: 27/04/2023 Aceptado: 22/12/2023

Análisis de datos

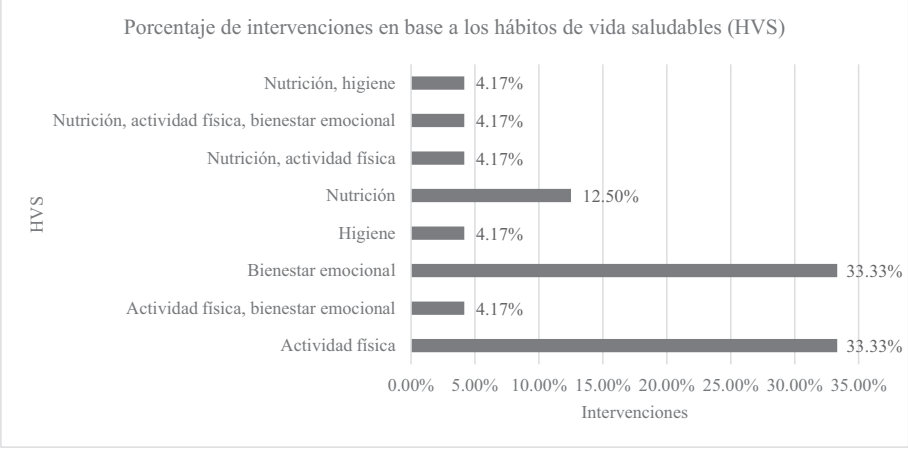
El análisis de datos se realizó mediante un análisis descriptivo y un análisis de contenido de los estudios. Las variables fueron agrupadas en distintas categorías: variables de los participantes (edad, curso), variables de contexto (fecha de publicación, escuelas de EP, y horario curricular), variables metodológicas (diseño de los estudios), variables extrínsecas (estudios publicados en revistas de impacto JCR), y variables de intervención (descripción, tipo de actividad, estructura, implicación familiar, formación, tareas, recursos).

Resultados y discusión

Análisis descriptivo

De los 24 artículos, el 37,93% incluyeron la actividad física (AF) como parte de su intervención, el 34,48% el bienestar emocional, el 20,69% la nutrición, y el 6,90% la higiene. Estos HVS fueron agrupados en algunas intervenciones, trabajando varias variables a la vez, por ello, hay intervenciones univariables y multivariables (Gráfico II).

GRÁFICO II. Número de intervenciones en base a los hábitos de vida saludables (HVS)

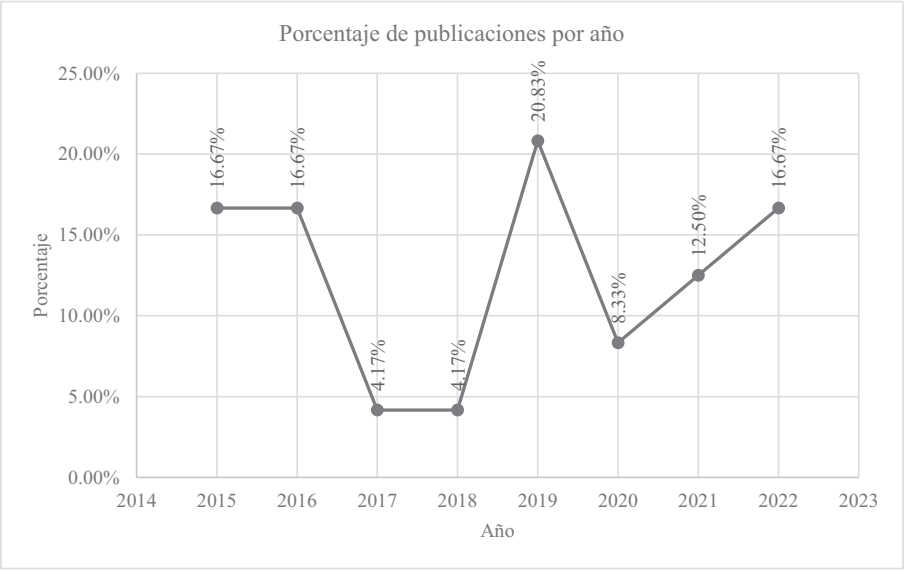


Fuente: elaboración propia.

Recibido: 27/04/2023 Aceptado: 22/12/2023

Estos artículos fueron publicados en 20 revistas distintas, siendo la revista *International Journal of Environmental Research and Public Health*, la más común (16,67%), después de *Trials* (8,33%). En cuanto a la fecha de publicación, el número de estudios publicados a lo largo de los siete años fue variando desde una publicación por año hasta 5, siendo 4 el número de mayor frecuencia (Gráfico III).

GRÁFICO III. Número de publicaciones por año



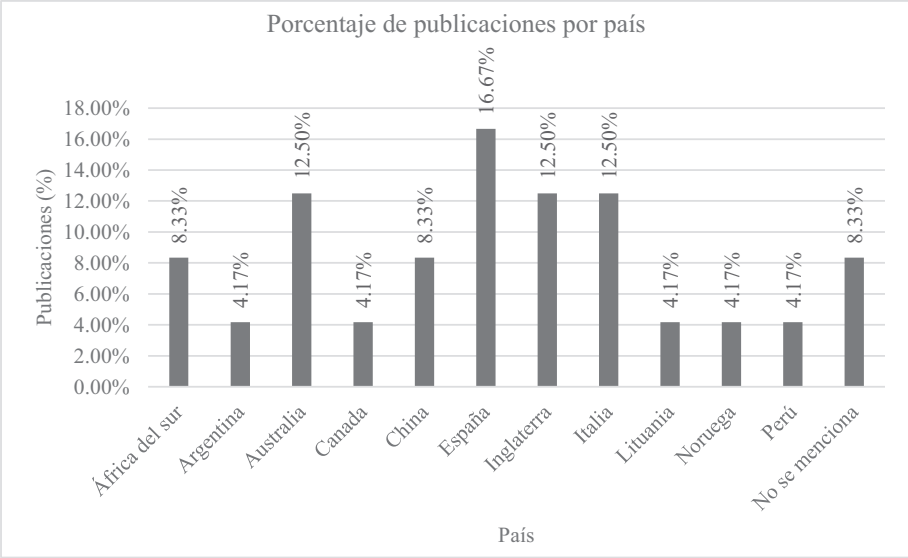
Fuente: elaboración propia.

Asimismo, se identificaron intervenciones realizadas en 12 países, siendo España el país que más intervenciones realizó (16,67%) seguido de Inglaterra (12,50%), Italia (12,50%) y Australia (12,50%) (Gráfico IV). En líneas generales la AF, el bienestar emocional y la nutrición fueron los HVS trabajados en estas intervenciones. La AF estuvo presente en un 53,85% de los estudios, el bienestar emocional en el 46,15%, y la nutrición en un 23,08%. Cabe mencionar que el 15,38% de estos estudios trabajaron más de un HVS, de los cuales todos fueron implementados en España. Este mismo país, España, fue el único país que trabajó los tres HVS mencionados (8,33% AF, 8,33% bienestar emocional y

Recibido: 27/04/2023 Aceptado: 22/12/2023

12,50% nutrición), mientras que Inglaterra (4,17% AF, 8,33% bienestar emocional y 0% nutrición), Italia (8,33% AF, 4,17% bienestar emocional y 0% nutrición) y Australia (8,33% AF, 4,17% bienestar emocional y 0% nutrición) no trabajaron la nutrición en ninguna de las intervenciones seleccionadas en este estudio.

GRÁFICO IV. Número de publicaciones por país



Fuente: elaboración propia.

(Continuada)

Los resultados generales de los 24 artículos (75% cuantitativos, y 25% mixtos) se recogen a continuación (Tabla IV).

TABLA IV. Objetivo y resultados generales de los estudios seleccionados

HVS	**Tipo de artículo**	**Artículo**	**Foco**	**Resultados generales**
Actividad física	cuantitativo	Atkinson et al. (2015)	un grupo	Aumenta el bienestar subjetivo de los niños, y la confianza en las interacciones sociales.
		Brustio et al. (2019)	quasi-experimental	Impacto positivo en las funciones físicas.
		Goswami et al. (2021)	quasi-experimental	Reducción del peso corporal y aumento de la masa muscular y ósea.
		Kliziene et al. (2021)	quasi-experimental	Impacto en la ansiedad somática, ansiedad de personalidad y ansiedad social. Aumento de los niveles de AF.
		Lu et al. (2022)	quasi-experimental	Mejora de la actitud del alumnado hacia la AF; y, la cognición conductual.
		Mavilidi et al. (2019)	quasi-experimental	Efectos positivos: habilidades de la función ejecutiva, y rendimiento de desplazamiento.
	mixto	Invernizzi et al. (2019)	quasi-experimental	Aumento del nivel de confianza, la competencia motriz, el disfrute y la cantidad de AF; tiempo en la dedicación a una tarea, y, satisfacción con las clases de EF y los estilos de enseñanza. Disminución de la pérdida de tiempo en EF.
		Watson et al. (2017)	quasi-experimental	Mejora de los resultados educativos clave.
Actividad física, bienestar emocional	mixto	Jimenez-Parra et al. (2022)	quasi-experimental	Mejora: funciones ejecutivas, rendimiento académico y aspectos psicosociales. Promueve la mejora de la calidad de vida relacionada con la salud.

Recibido: 27/04/2023 Aceptado: 22/12/2023

TABLA IV. Objetivo y resultados generales de los estudios seleccionados *(Continuada)*

HVS	Tipo de artículo	Artículo	Foco	Resultados generales
Bienestar emocional	cuantitativo	Crescentini et al. (2016)	quasi-experimental	Mejora: capacidades cognitivas, emocionales y sociales. Efectos positivos: habilidades atencionales, síntomas del TDAH y funciones emocionales.
		Kishida et al. (2022)	un grupo	Beneficios: en niveles de apoyo entre iguales y la soledad de los estudiantes; y, la autoeficacia de los padres.
		Schonnert et al. (2015)	experimental	Mejora: funciones ejecutivas, medidas de bienestar autodeclaradas, comportamiento prosocial autodeclarado y declarado por los compañeros; y, rendimiento en matemáticas.
		Shum et al. (2019)	quasi-experimental	Mejora; conocimientos sobre salud mental.
		Sitzer et al. (2015)	un grupo	Mejora: resiliencia y funcionamiento social y emocional.
		Suldo et al. (2015)	un grupo	Mejora: emociones positivas, satisfacción global con la vida y satisfacción consigo mismos, con sus amigos y con el entorno en el que vivían.
		Vickery et al. (2016)	quasi-experimental	Mejora: metacognición. Reduce: afecto negativo.
	mixto	Humphrey et al. (2018)	ensayo controlado aleatorizado	Mejora: habilidades sociales de los niños; percepción del apoyo social y de sus compañeros, Reduce: exclusiones Impacto significativo: bienestar psicológico de los niños.
Higiene	mixto	Oberto et al. (2020)	un grupo	Mejora: retroalimentación de saberes y la construcción colectiva de conocimientos; y, creación de un entorno favorable para promover comportamientos saludables.

(Continuada)

Recibido: 27/04/2023 Aceptado: 22/12/2023

TABLA IV. Objetivo y resultados generales de los estudios seleccionados *(Continuada)*

HVS	Tipo de artículo	Artículo	Foco	Resultados generales
Nutrición	cuantitativo	de Villiers et al. (2016)	quasi-experimental	Mejora: conocimientos sobre nutrición y la autoeficacia de los niños.
		Martinez-Garcia y Trescastro-López et al. (2016)	un grupo	Mejora: conocimiento y consumo de alimentos saludables. Disminución: consumo de alimentos poco saludables.
		García y Fernández (2022)	quasi-experimental	Mejora: adherencia a la dieta mediterránea.
Nutrición, actividad física	cuantitativo	Gerber et al. (2020)	Ensayo aleatorizado, controlado con placebo	(1) Comprensión de los efectos de una intervención, que combina la AF y la suplementación con múltiples micronutrientes. (2) Nuevos conocimientos sobre la capacitación de las escuelas para implementar datos de AF. Actualización del estado de las deficiencias de micronutrientes, los marcadores de riesgo inflamatorio y cardiovascular y las enfermedades transmisibles en las comunidades seleccionadas.
Nutrición, actividad física, bienestar emocional	mixto	Santos-Beneit et al. (2019)	un grupo	Incremento: puntuaciones KAH hacia un estilo de vida saludable.
Nutrición, higiene	cuantitativo	López et al. (2021)	quasi-experimental	Mejora: HVS, consumo de agua, frutas y verduras; y, prácticas de lavado de manos e higiene dental. Disminución: consumo de refrescos.

Fuente: elaboración propia.
Nota: TDAH: trastorno por déficit de atención e hiperactividad; KAH: Knowledge, Attitudes, and Habits; EF: educación física.

Recibido: 27/04/2023 Aceptado: 22/12/2023

Análisis de contenido de los estudios

El análisis de contenido se dividió en dos secciones. Primero, se analizaron los contenidos generales de los estudios en base a 3 ejes (Tabla V): población, asignatura, y agente que dirige la intervención. En segundo lugar, se analizaron con más profundidad las intervenciones llevadas a cabo, esta vez en 7 ejes (Tabla VI): descripción; estructura de la intervención: frecuencia (F), tiempo (T), intensidad (I); tipo de actividades; formación; implicación de la familia; tareas adicionales; y recursos.

TABLA V. Resumen del análisis de los contenidos generales de los estudios

Artículo	**Población**	**Asignatura**	**Persona que lleva a cabo la intervención**
Atkinson et al. (2015)	$N_{alumnado}$ = 28 7 y 8 años	IC: AC	Bailarines, PA
Brustio et al. (2019)	$N_{alumnado}$ = 795 6-9 años	IC: AC (no EF)	PA
Crescentini et al. (2016)	$N_{alumnado}$ = 31 7-8 años	IC: AC	Instructores de meditación mindfulness
de Villiers et al. (2016)	$N_{alumnado}$ = 998 4º curso	IC: AC	Educadores
Gerber et al. (2020)	$N_{alumnado}$ = 330 6-12 años	IC: AC	PA, Instructores
Goswami et al. (2021)	$N_{alumnado}$ = 917 1-6º curso	IC: AC	PA
Humphrey et al. (2018)	$N_{alumnado}$ = 5218 7-9 años 3-5º cursos	IC: AC	PA
Invernizzi et al. (2019)	$N_{alumnado\text{-}cuantitativo}$ = 121 $N_{alumnado\text{-}cualitativo}$ = 28 10-11 años	AC: EF	Estudiantes graduados de EF, PA
Jimenez-Parra et al. (2022)	$N_{alumnado}$ = 250	IC: AC	PA
Kishida et al. (2022)	$N_{alumnado}$ = 37 10 y 12 años	IC: AC	PA, familia
Kliziene et al. (2021)	$N_{alumnado}$ = 364 6-9 años	AC: EF	Profesorado de EF

(Continuada)

TABLA V. Resumen del análisis de los contenidos generales de los estudios *(Continuada)*

Artículo	**Población**	**Asignatura**	**Persona que lleva a cabo la intervención**
López et al. (2021)	$N_{alumnado}$ = 108 6-13 años	IC: AC	Estudiante de enfermería, PA
Lu et al. (2022)	$N_{alumnado}$ = 366 8-11 años	AC: EF	Profesorado de EF, familia
Martinez-Garcia y Trescastro-López (2016)	$N_{alumnado}$ = 28 8 y 9 años 3º curso	IC: AC	Dietista-nutricionista, PA, profesorado tutor/a, profesorado de EF
Mavilidi et al. (2019)	$N_{alumnado}$ = 10 escuelas 4º curso	AC: inglés	PA
Oberto et al. (2020)	$N_{alumnado}$ = 564 4-10 años 1-3º curso	AC: ciencias sociales	Investigadores
García y Fernández. (2022)	$N_{alumnado}$ = 38 6-7 años 1º curso	AC: EF	NSE
Santos-Beneit et al. (2019)	$N_{alumnado}$ = 1770 1-6º curso	IC: AC	NSE
Schonnert et al. (2015)	$N_{alumnado}$ = 99 9-12 años 4-5º curso	IC: AC	NSE
Shum et al. (2019)	$N_{alumnado}$ = 459 8-12 años	IC: AC	Licenciados o estudiantes de un máster en Asesoramiento o Asesoramiento y Psicología Clínica o profesores con formación previa
Sitzer et al. (2015)	$N_{alumnado}$ = 43 9-12 años 4-6º curso	AC: grupo único de asignaturas que promueven las habilidades de afrontamiento y resiliencia	NSE
Suldo et al. (2015)	$N_{alumnado}$ = 12 9-10 años 4º curso	AC: matemática, ciencias	PA, psicólogo escolar e Investigadores
Vickery et al. (2016)	$N_{alumnado}$ = 71 7-9 años	AC: educación personal y social	PA

Recibido: 27/04/2023 Aceptado: 22/12/2023

TABLA V. Resumen del análisis de los contenidos generales de los estudios *(Continuada)*

Artículo	**Población**	**Asignatura**	**Persona que lleva a cabo la intervención**
Watson et al. (2017)	$N_{escuelas}$ = 6 8-10 años 3-4º curso	IC: AC	PA, investigadores

Fuente: elaboración propia.

Nota: actividad física; : bienestar emocional; : nutrición; : higiene; EF: educación física; AC: asignatura curricular; IC: integración curricular; NSE: no se especifica; PA: profesorado de aula

Población

De los 24 estudios, 98,84% indicaron el curso o la edad del alumnado. Las intervenciones que informaban sobre el curso fueron clasificados según la edad: primero (6-7 años), segundo (7-8 años), tercero (8-9 años), cuarto (9-10 años), quinto (10-11 años), y, sexto (11-12 años).

La edad es relevante para educar al alumnado en HVS (García-Muñoz et al., 2022; Langford et al., 2014). En este estudio, las edades de nueve (22%), diez (18%) y ocho (17%) años manifiestan una mayor congregación de intervenciones sobre los HVS en el currículum, y aunque en revisiones previas la edad no reporta un impacto en las intervenciones (Hung et al., 2015), hay evidencia de que el alumnado mayor muestra un mejor entendimiento de los conceptos de la intervención (Stice et al., 2006).

Asignatura

Todos los estudios mencionaron las asignaturas donde se llevaron a cabo las intervenciones.

Todas formaron parte del currículo, y es que el horario lectivo es un periodo útil para educar y adquirir los HVS (Hahnraths et al., 2021; Pulido-Gil et al., 2022). El 62,53% fueron realizadas en integraciones curriculares (inclusión de más de una asignatura curricular en el estudio); y, el resto en una única asignatura curricular. Todas las integraciones curriculares fueron dirigidas en asignaturas curriculares no especificadas (58,33%), y en una

de ellas la educación física (EF) fue excluida de la integración (4,17%). En cuanto a las asignaturas curriculares, la EF (16,67%) fue la asignatura donde se dirigieron más intervenciones. La EF tiene la capacidad de alcanzar los HVS y generar un impacto relevante en ellos (Invernizzi et al., 2019; Kliziene et al., 2021). Aun así, es importante intervenir en el resto de las asignaturas, ya que muchas se imparten en gran medida de manera sedentaria (Daly-Smith et al., 2021), pudiendo convertirse en una estrategia para fomentar y adquirir los HVS a través de los planes curriculares. En el Estado Español, por ejemplo, los HVS están estrechamente relacionados con el área de Conocimientos del Medio Natural, Social y Cultural (Real Decreto 157/2022). En este estudio, el resto de asignaturas representan un 4,17% de las intervenciones cada una de ellas.

Agente

De los 24 estudios, 94,59% aportan información sobre el agente. Algunas intervenciones se desarrollaron con la ayuda de varios miembros.

El profesorado es el agente encargado de impartir la programación de aula y el proceso de aprendizaje del alumnado (Real Decreto 157/2022). Así, la revisión indica que el profesorado de las escuelas fue el encargado principal del desarrollo de las intervenciones (51,35%). Concretamente, el profesorado de aula fue quien más intervenciones llevó a cabo (43,24%), mientras que el resto fueron dirigidos por el de EF (5,41%) o por profesores formados ad hoc (2,70%). Esto es un dato de interés, ya que las intervenciones llevadas a cabo por parte del profesorado escolar tienen mayor probabilidad de integrarse en los planes de estudios y de mantenerse a lo largo del tiempo (de Bourdeaudhuij et al., 2010; Fairclough et al., 2013). Además, cabe destacar la baja incidencia del profesorado de EF ya que estudios previos han demostrado que estos tienen un rol fundamental en experiencias relacionadas con la salud (Invernizzi et al., 2019; López-Gil, et al. 2020). Finalmente, y en contraposición a estudios previos (Hung et al., 2015; Stice eta l., 2006) las intervenciones dirigidas a la promoción de las HVS parecen reservar un espacio formativo alto (44%) a los agentes externos.

En la Tabla VI, se analizan con profundidad las intervenciones, a saber: estructura, tipo de contenido, tipo de actividad, implicación de la familia, tareas adicionales y recursos.

Recibido: 27/04/2023 Aceptado: 22/12/2023

TABLA VI. Resumen del análisis de los contenidos de las intervenciones de los estudios

Artículo	Estructura	Tipo de contenido	Tipo de actividades	Formación	Implicación de la familia	Tareas adicionales	Recursos
Atkinson et al. (2015)	F: no T: no D: 3 meses I: no	Procedimental Actitudinal Conceptual	Movimiento corporal (EC, HM) Concienciación Valores y emociones Adquisición de conocimiento	✗	✓	✗	Audiovisual
Brustio et al. (2019)	F: 3 veces a la semana T: 15 min D: 3 meses I: no	Procedimental	Movimiento corporal (HM)	✓	✗	✗	NSE
Crescentini et al. (2016)	F: 3 veces a la semana T: 1-2 semanas (30min); 3-4 semanas (45-55min); 7-8 semanas (1h 30min) D: 8 semanas I: no	Procedimental Actitudinal	Relajación y atención Concienciación	✗	✗	✗	Audiovisual
de Villiers et al. (2016)	F: no T: no D: 3 años I: no	Procedimental Actitudinal Conceptual	Adquisición de conocimiento Concienciación Experimental-manipulativo	✓	✓	✗	Impresos
Gerber et al. (2020)	F: no T: no D: no I: no	Procedimental	Movimiento corporal (DA, EF, EC)	✓	✓	✗	NSE
Goswami et al. (2021)	F: no T: 135 min D: no I: moderada-vigorosa	Procedimental	Movimiento corporal (HM)	✗	✗	✗	NSE

(Continuada)

TABLA VI. Resumen del análisis de los contenidos de las intervenciones de los estudios *(Continuada)*

Artículo	Estructura	Tipo de contenido	Tipo de actividades	Formación	Implicación de la familia	Tareas adicionales	Recursos
Humphrey et al. (2018) 	F: 2 vecer por semana T: 30-40 min D: 1 año académico I: no	Conceptual Actitudinal	Adquisición de conocimiento Concienciación	✓	✓	✓	Impresos
Invernizzi et al. (2019) 	F: 2h a la semana T: 2h D: 12 semanas I: no	Procedimental	Valores y emociones	✓	×	×	NSE
Jimenez-Parra et al. (2022) 	F: 2-4 veces al día (10-16 sesiones a la semana) T: 1h D: 1 año académico I: vigorosa	Actitudinal Procedimental	Valores y emociones Concienciación Movimiento corporal (DA)	✓	×	✓	Audiovisual
Kishida et al. (2022) 	F: no T: 40-60 min D: 12 meses I: no	Actitudinal	Valores y emociones	×	✓	×	NSE
Kliziene et al. (2021) 	F: 3 veces a la semana T: 30-45 min D: 8 mese I: no	Procedimental	Movimiento corporal (HM) Relajación y atención	×	×	×	Impresos
López et al. (2021) 	F: no T: 1h D: 6 meses I: no	Actitudinal Procedimental	Concienciación Experimental-manipulativo	✓	×	×	NSE

Artículo	Estructura	Tipo de contenido	Tipo de actividades	Formación	Implicación de la familia	Tareas adicionales	Recursos
Lu et al. (2022)	F: 1 vez a la semana (hojas EMWKG); 3 veces a la semana (deberes de EF) T: no D: 15 semanas I: no	Procedimental Actitudinal Conceptual	Adquisición de conocimiento Movilidad corporal (HM) Experimental-manipulativo	✗	✓	✓	Impresos, ilustraciones; audiovisual
Martinez-Garcia y Trescastro-López (2016)	F: no T: 60-90 min D: 1 mes (5 sesiones) I: no	Conceptual Procedimental	Adquisición de conocimiento Experimental-manipulativo	✗	✓	✗	NSE
Mavilidi et al. (2019)	F: 3 veces a la semana T: 40 min D: 6 semanas I: no	Procedimental	Movimiento corporal (HM)	✓	✗	✗	NSE
Oberto et al. (2020)	F: no T: no D: 3 años (17 talleres) I: no	Actitudinal Procedimental	Concienciación Experimental-manipulativo	✗	✗	✗	Libro
García y Fernández (2022)	F: 1 vez a la semana T: no D: 5 semanas I: no	Procedimental Actitudinal Conceptual	Experimental-manipulativo Concienciación	✓	✗	✗	NSE
Santos-Beneit et al. (2019)	F: no T: no D: no I: no	Procedimental Actitudinal Conceptual	Adquisición de conocimiento Concienciación Valores y emociones	✓	✓	✓	Audiovisuales, manipulativos

(Continuada)

TABLA VI. Resumen del análisis de los contenidos de las intervenciones de los estudios *(Continuada)*

Artículo	**Estructura**	**Tipo de contenido**	**Tipo de actividades**	**Formación**	**Implicación de la familia**	**Tareas adicionales**	**Recursos**
Schonnert et al. (2015)	F: 1 vez a la semana (lección) / 3 veces al día (prácticas básicas de mindfulness) T: 40-50 min D: 4 meses (12 sesiones) I: no	Procedimental	Valores y emociones	×	×	×	Audiovisual
Shum et al. (2019)	F: no T: 20 min D: 11 lecciones digitales basadas en juegos, y, 8 enseñanzas en clase. I: no	×	Valores y emociones	✓	×	×	Digital
Sitzer et al. (2015)	F: 1 vez a la semana T: 60 min D: 14 semanas I: no	Procedimental conceptual	Adquisición de conocimiento Relajación y atención Valores y emociones (mindfulness) Experimental-manipulativo Concienciación	×	×	×	Impresos ilustraciones
Suldo et al. (2015)	F: no T: 30-50 min D: 11 sesiones I: no	Procedimental Conceptual	Valores y emociones Adquisición de conocimiento	×	×	×	NSE

Artículo	Estructura	Tipo de contenido	Tipo de actividades	Formación	Implicación de la familia	Tareas adicionales	Recursos
Vickery et al. (2016)	F: 1 tema por semana (excepto 2 temas que llevaron 2 semanas) T: 30-60 min D: 12 clases de media hora en total I: no	Procedimental	Relajación y atención (mindfulness)	✓	✓	✓	NSE
Watson et al. (2017)	F: 3 veces al día T: 5 min D: 6 semanas I: moderada	Procedimental	Movimiento corporal (DA, EC, HM)	✓	✓	✗	Audiovisual

Fuente: elaboración propia

Nota: : actividad física; : bienestar emocional; : nutrición; : higiene; ✓: si; ✗: no; F: frecuencia; T: tiempo; D: duración; I: intensidad; HM: habilidad motriz; DA: descanso activos; EC: expresión corporal; NSE: no se especifica.

Estructura

La estructura de las intervenciones fue clasificada según la frecuencia (F), tiempo (T), duración (D), e intensidad (I). La frecuencia hizo referencia a la cantidad de veces que se realizó la intervención. Se observó una mayor ponderación de las intervenciones realizadas una y tres veces por semana (20%). Esto podría deberse a que los hábitos requieren de repetición, y aprendizaje (Mosocoso-Loaiza y Días-Heredia, 2018; Oberto et al., 2020). En lo que se refiere al tiempo que duraban las actividades de las intervenciones, las de 30-60 minutos (44%) fueron las más comunes. El periodo de las intervenciones indicaba la duración de la misma (sesiones, semanas, o años). La duración es un importante determinante de la eficacia, pero se desconoce la idoneidad de la misma, subrayando la necesidad de investigar más al respecto (Amini et al., 2015). En este caso, el 16% de los estudios señalaron realizar 8 y 12 sesiones, sin indicar el periodo exacto de las intervenciones. Las intervenciones de entre 5-8 semanas mostraron mayor frecuencia (16%) seguido de las de 9-12 y 13-16 semanas, ambas con un 12%. Por último, pocas intervenciones aludían a la intensidad realizada en las actividades (8%), siendo las actividades moderadas (4%) y vigorosas (4%) las más corrientes. Esto puede deberse a que la OMS sugiere incorporar actividades de intensidad moderada y vigorosa en la población infantil y adolescente (5-17 años) (OMS, 2020).

Tipo de contenido y actividad

El 98,84% de los estudios indicaron el tipo de contenido, siendo el procedimental el contenido más habitual (49%) después de la actitudinal (26%) y la conceptual (23%). Las actividades que requerían concienciación fueron las más utilizadas (22%), después de las de movimiento corporal (20%), y adquisición de conocimiento, y las de los valores y emociones, las cuales reportaron un 17%. Dentro de las actividades de movimiento corporal se exponía una mayor frecuencia de las habilidades motrices (46%), la expresión corporal (23%) y el descanso activo (23%). Algunas intervenciones indicaron más de un tipo de contenido o tipo de actividad, reportando una mayor frecuencia.

Recibido: 27/04/2023 Aceptado: 22/12/2023

Formación, implicación de la familia, y tareas adicionales

La mitad de los estudios ofrecieron una formación previa a la intervención, resaltando la importancia de formar a los diferentes agentes implicados (Mcloughlin et al., 2020).

La implicación de la familia es fundamental (Hung et al., 2015; Oberto et al., 2020). Así, no sorprende el hecho de que casi el 42% de los estudios implicasen a la familia. Además, las investigaciones sobre AF y nutrición previas indican la eficacia de las intervenciones escolares al reforzarlos con un componente de intervención familiar (Hung et al., 2015; Oberto et al., 2020).

Por último, pocas intervenciones añadieron tareas opcionales, 20,83%, aunque autores sugieren la importancia de añadir tareas en casa para reforzar la participación familiar (McCullog et al., 2019).

Recursos

En esta revisión, el 54,17% de los estudios indicaron la utilización de algún recurso. Concretamente, el 25% recursos audiovisuales, el 17,86% impresos y el 7,14% ilustraciones. Resulta llamativa la baja influencia del uso de la tecnología, ya que estudios recientes como el de Langarizadeh et al. (2021) muestran que la tecnología supone una herramienta fundamental para la mejora de las intervenciones de promoción de la salud.

Conclusiones

Esta revisión sistemática tuvo como objetivo, analizar las intervenciones realizadas en el ámbito curricular de EP para mejorar los HVS en el alumnado.

La revisión de la literatura permite confirmar que no existe un marco declarativo sobre las dimensiones de los HVS en EP. Así, este trabajo ofrece una visión indiciaria sobre la situación conceptual y teórica de dicho concepto, poniendo a la luz las diferentes dimensiones a través de las cuales está estructurándose el mismo: AF, nutrición, bienestar emocional e higiene. Estos resultados, por un lado, pusieron de relieve las dimensiones más utilizadas en el ámbito curricular de EP para mejorar los

HVS del alumnado. Parece ser que tanto la AF como el bienestar emocional fueron las dimensiones pilares en las intervenciones curriculares de EP. Por otro lado, estos resultados evidenciaron la existencia de una carencia sobre la categorización de los HVS, imprescindibles para la salud y bienestar presente y futura de la población.

Los estudios analizados pusieron de manifiesto la preponderancia de la asignatura de EF para mejorar los HVS en el ámbito curricular. La EF siempre ha estado ligada de alguna manera a la salud, fundamentalmente, por la promoción del movimiento que a esta se le ha atribuido. Aun así, esta revisión mostró que los HVS también pueden ser desarrollados y fomentados en otros ámbitos curriculares, como pueden ser las ciencias sociales, el inglés o las matemáticas. Fue de gran importancia resaltar dicha contribución, teniendo en cuenta la relación existente entre la salud y bienestar y la educación. El alumnado pasa gran parte de su día en la escuela en espacios curriculares, por lo que sería enriquecedor utilizar estos y otros (ciencias naturales, áreas artísticas, etcétera) para fomentar los HVS.

En cuanto a los contenidos utilizados en las intervenciones, a pesar de la consideración que tienen los contenidos conceptuales, ya que ofrecen la posibilidad de conocer más de cerca ciertos conceptos relacionados con los HVS mencionados, éstas fueron escasas, y, se pusieron de relieve los contenidos procedimentales (dando importancia tanto a la ejecución de acciones como a la consecución de metas) y los actitudinales (valorando la sensibilización existente en las conductas saludables). De este estudio destaca la importancia de las actividades de concienciación; y es que los HVS requieren de sensibilización por parte del estudiantado para integrar de manera rutinaria las funciones tan básicas. El alumnado debería de ser consciente de los beneficios y las consecuencias que los HVS pueden generar a futuro.

En cuanto a la ejecución de las intervenciones y las personas participantes en ellas, cabe destacar que la mayoría de las intervenciones analizadas incluyeron una formación previa y la presencia de la familia. Parece evidente que la mejora de los HVS mediante intervenciones requería de una formación previa, poniendo de manifiesto la importancia y responsabilidad del agente encargado de la implementación, así como el de la familia.

Finalmente, en base a los resultados, se pudo obtener una visión sobre el desarrollo y la implementación de los HVS en el plano educativo. Así,

Recibido: 27/04/2023 Aceptado: 22/12/2023

sería de gran interés que las futuras líneas de investigación considerasen las limitaciones de este estudio, para así, poder avanzar en la investigación de sobre los HVS en el currículum de EP. Por un lado, el hecho de atender al resto de agentes de la educación, especialmente, al profesorado y a la familia, ayudaría a tener una visión más global de las intervenciones que se deben de implementar en las aulas. Por otro lado, el análisis de contenido específico de los tipos de actividades y contenidos, recursos utilizados, instrumentos de medida, tipo de formación y tareas adicionales, enriquecería la línea de investigación mencionada. Se obtendría una visión más detallada de las intervenciones que ayudan a mejorar y adquirir los HVS en el ámbito curricular. Por ello, de cara a líneas de investigación futuras cabe plantear el análisis en profundidad de las intervenciones, la elaboración del grado de consecución del objetivo, y la comparación de las variables analizadas con cada HVS.

Agradecimientos y financiación

Este estudio ha sido financiado por la Universidad de Deusto (Ref: FPI UD_2022_10).

Referencias bibliográficas

Akseer, N., Mehta, S., Wigle, J., Chera, R., Brickman, Z. J., Al-Gashm, S., Sorichetti, B., Vandermorris, A., Hipgrave, D. B., Schwalbe, N., y Bhutta, Z. A. (2020). Non-communicable diseases among adolescents: current status, determinants, interventions and policies. *BMC Public Health 20*(1908), 1-20. https://doi.org/10.1186/s12889-020-09988-5

Amini M, Djazayery A, Majdzadeh R, Taghdisi M-H, & Jazayeri S. (2015). Effect of Schoolbased Interventions to Control Childhood Obesity: A Review of Reviews. *International Journal of Preventive Medicine*, *6*(68), 1-15. https://doi.org/10.4103/2008-7802.162059

Atkinson, S., & Scott, K. (2015). Stable and destabilised states of subjective well-being: dance and movement as catalysts of transition. *Social & Cultural Geography, 16*(1), 75-94. https://doi.org/10.1080/14649365.2014.950689

Bonell, C., Parry, W., Wells, H., Jamal, F., Fletcher, A., Harden, A., Thomas, J., Campbell, R., Petticrew, M., Murphy, S., Whitehead, M., & Moore, L. (2013). The effects of the school environment on student health: A systematic review of multi-level studies. *Health & Place; Place, 21*, 180-191. https://doi.org/10.1016/j.healthplace.2012.12.001

Brustio, P. R., Mulasso, A., Marasso, D., Ruffa., Ballatore, a., Moisè, P., Lupo, C., Rainoldi, A., & Boccia, G. (2019). The Daily Mile: 15 Minutes Running Improves the Physical Fitness of Italian Primary School Children. *International Journal of Environmental Research and Public Health, 16*(3921), 1-10. https://doi.org/10.3390/ijerph16203921

Clark, H., Coll-Seck, A. M., Banerjee, A., Peterson, S., Dalglish, S. L., Ameratunga, S., Balabanova, D., Bhan, M. K., Bhutta, Z. A., Borrazzo, J., Claeson, M., Doherty, T., El-Jardali, F., George, A. S., Gichaga, A., Gram, L., Hipgrave, D. B., Kwamie, A., Meng, Q., & Costello, A. (2020). A future for the world's children? A WHO–UNICEF–Lancet Commission. *The Lancet, 395*, 605-658. https://doi.org/10.1016/S0140-6736(19)32540-1

Crescentini, C., Capurso, V., Furlan, S., & Fabbro, F. (2016). Mindfulness-Oriented Meditation for Primary School Children: Effects on Attention and Psychological Well-Being. *Frontiers Psychology, 7*(805), 1-12. https://doi.org/10.3389/fpsyg.2016.00805

Daly-Smith, A., Hobbs, M., Morris, J. L., Defeyter, M. A., Resaland, G. K., & McKenna, J. (2021). Moderate-to-Vigorous Physical Activity in Primary School Children: Inactive Lessons Are Dominated by Maths and English. *International Journal of Environmental Research and Public Health, 18*(990), 1-14. https://doi.org/10.3390/ijerph

de Bourdeaudhuij, I., Cauwenberghe, E. V., Spittaels, H., Oppert, J. M., Rostami, C., Brug, J., Lenthe, F. V., Lobstein, D. & Maes, L. (2010). School-based interventions promoting both physical activity and healthy eating in Europe: a systematic review within the HOPE project. *Obesity reviews, 12,* 205-2016. https://doi.org/10.1111/j.1467-789X.2009.00711.x

de Villiers, A., Steyn, N. P., Draper, C. E., Hill, J., Gwebushe, N., Lambert, E. V., & Lombard, C. (2016). Primary School Children's Nutrition Knowledge, Self-Efficacy, and Behavior, after a Three-Year Healthy Lifestyle Intervention (HealthKick). *Ethnicity & Disease, 26*(2), 171-180. https://doi.org/10.18865/ed.26.2.171

Recibido: 27/04/2023 Aceptado: 22/12/2023

Fairclough, S. J., Hackett, A. F., Davies, I. G., Gobbi, R., Mackintosh, K. A., Warburton, G. L., Stratton, G., van Sluijs, E., & Boddy, L. (2013). Promoting healthy weight in primary school children through physical activity and nutrition education: a pragmatic evaluation of the CHANGE! randomised intervention study. *BMC Public Health, 13*(626), 1-14. http://www.biomedcentral.com/1471-2458/13/626

Gámez-Calvo, L., Hernández-Beltrán, V., Pimienta-Sánchez, L. P., Delgado-Gil, S., & Gamonales, J. M. (2022). Revisión sistemática de programas de intervención para promover hábitos saludables de actividad física y nutrición en escolares españoles. *Revista de la Sociedad Latinoamericana de Nutrición, 72*(4), 294-305. https://doi.org/10.37527/2022.72.4.007

García, E., & Fernández, G. (2022). Intervención educativa mediante una propuesta de gamificación para mejorar la adhesión a la dieta mediterránea en estudiantes gallegos de primaria. *Retos, 44,* 128-135. https://doi.org/10.47197/retos.v44i0.90142

García-Muñoz, S., Barlińska, J., Wojtkowska, K., Da Quinta, N., Baranda, A., Alfaro, B., & Santa Cruz, E. (2022). Is it possible to improve healthy food habits in schoolchildren? A cross cultural study among Spain and Poland. *Food Quality and Preference, 99*, 104534. https://doi.org/10.1016/j.foodqual.2022.104534

Gerber, M., Ayekoé, S.A., Beckmann1, J., Bonfoh, B., Coulibaly, J.T., Daouda, D., du Randt, R., Finda, L., Gall, S., Mollel, G.J., Langl, C., Long, K.Z., Ludygal, S., Masanja, H., Müller, I., Nqweniso, S., Okumu, F., Probst-Hensch, N., Pühsel, U., & Utzinger, J. (2020). Effects of school-based physical activity and multi-micronutrient supplementation intervention on growth, health and wellbeing of schoolchildren in three African countries: the KaziAfya cluster randomised controlled trial protocol with a 2 × 2 factorial design. *Trials 21*(22), 1-17. https://doi.org/10.1186/s13063-019-3883-5

Goswami, N., Trozic, I., Valand, M., & Morten, P. (2021). The effect of physical activity intervention and nutritional habits on anthropometric measures in elementary school children: the health oriented pedagogical project (HOPP). *International Journal of Obesity, 45*, 1677-1686. https://doi.org/10.1038/s41366-021-00830-5

Hahnraths, M. T. H., Willeboordse, M., Jungbauer, A. D. H. M., de Gier, C., Schouten, C., & van Schayck, C. P. (2021). Mummy, Can I Join a Sports Club? A Qualitative Study on the Impact of Health-Promoting Schools

on Health Behaviours in the Home Setting. *International Journal of Environmental Research and Public Health, 18*(22), 12219. https://doi.org/10.3390/ijerph182212219

Harris, J., & Cale, L. (2018). *Promoting active lifestyles in schools*. Champaign, IL: Human Kinetics.

Humphrey, N., Hennessey, A., Lendrum, A., Wigelsworth, M., Turner, A., Panayiotou, M., Joyce, C., Pert, K., Stephens, E., Wo, L., Squires, G., Woods, K., Harrison, M., & Calam, R. (2018). The PATHS curriculum for promoting social and emotional well-being among children aged 7–9 years: a cluster RCT. *Journal of Public Health Research, 6*(10), 1-142. https://doi.org/10.3310/phr06100

Hung, L. S., Tidwell, D., Hall, M. E., Lee, M., Briley, C. A., & Hunt, B. P. (2015). A meta-analysis of school-based obesity prevention programs demonstrates limited efficacy of decreasing childhood obesity. *Nutrition Research, 35,* 229-240. http://dx.doi.org/10.1016/j.nutres.2015.01.002

Invernizzi, P., Crotti, M., Bosio, A., Cavaggioni, L., Alberti, G., & Scurati, R. (2019). Multi-Teaching Styles Approach and Active Reflection: Effectiveness in Improving Fitness Level, Motor Competence, Enjoyment, Amount of Physical Activity, and Effects on the Perception of Physical Education Lessons in Primary School Children. *Sustainability, 11*(405), 1-20. https://doi.org/10.3390/su11020405

Jiménez-Parra, J. F., Belando-Pedreño, N., López-Fernández, J., García-Vélez, A. J., & Valero-Valenzuela, A. (2022). "ACTIVE VALUES": An Interdisciplinary Educational Programme to Promote Healthy Lifestyles and Encourage Education in Values—A Rationale and Protocol Study. *Applied Science, 12*(8073), 1-20. https://doi.org/10.3390/app12168073

Jones, M., Defever, E., Letsinger, A., Steele, J., & Mackintosh, K. A. (2020). A mixed-studies systematic review and meta-analysis of school-based interventions to promote physical activity and/or reduce sedentary time in children. *Journal of Sport and Health Science, 9*(1), 3-17. https://doi.org/10.1016/j.jshs.2019.06.009

Kishida, Y., Brennan-Jones, C. g., Runions, K., Vithiatharan, R., Hancock, K., Brown, M., Eikelboom, R. H., Coffin, J., Kickett-Tucker, C., Li, I. W., Epstein, M., Falconer, S. E., & Cross, D. (2022). Supporting the Social–Emotional Well-Being of Elementary School Students Who Are Deaf and Hard of Hearing: A Pilot Study. *American Speech-Language-Hearing Association,* 1-14. https://doi.org/10.1044/2022_LSHSS-21-00178

Recibido: 27/04/2023 Aceptado: 22/12/2023

Kliziene, I., Cizauskas, G., Sipaviciene, S., Aleksandraviciene, R., & Zaicenkoviene, K. (2021). Effects of a Physical Education Program on Physical Activity and Emotional Well-Being among Primary School Children. *International Journal of Environmental Research and Public Health, 18*(14), 7536. https://doi.org/10.3390/ijerph18147536

Langarizadeh, M., Sadeghi, M., As'habi, A., Rahmati, P., & Sheikhtaheri, A. (2021). Mobile apps for weight management in children and adolescents; An updated systematic review. *Patient Education and Counseling, 104*(9), 2181-2188. https://doi.org/10.1016/j.pec.2021.01.035

Langford, R., Bonell, C. P., Jones, H. E., Pouliou, T., Murphy, S. M., Waters, E., Komro, K. A., Gibbs, L. F., Magnus, D., Campbell, R., & Langford, R. (2014). The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *Cochrane Database of Systematic Reviews, 2014*(4), CD008958. https://doi.org/10.1002/14651858.CD008958.pub2

López, M., Alcoceba, I., Castro, M. J., Cao, M. J., García, S., Frutos, M., & Jiménez, J. M. (2021). Assessment of an Educational Intervention to Improve Healthy Life Habits in Children Living in Vulnerable Socioeconomic Conditions. *International Journal of Environmental Research and Public Health, 18*(4495), 1-11. https://doi.org/10.3390/ijerph18094495

López-Gil, J. F., Renato, F., & Yuste, J. L. (2020). RETOS. Nuevas tendencias en Educación Física, Deporte y Recreación Número 37, 2020 (1º semestre)- 786 -Programas de intervención para la promoción de hábitos alimenticios saludables en escolares españoles practicantes de Educación Física: una revisión sistemática. *Retos, 37,* 786-792. https://doi.org/10.47197/retos.v37i37.69931

Lu, Y., Yu, K., Jin, J., & Gan, X. (2022). The Effects of a Multicomponent Social Support Intervention on Physical Fitness and Exercise Attitude in Children: A 12-Week Randomized Controlled Trial. *International Journal of environmental Research and Public Health, 19*(9922)1-19. https://doi.org/10.3390/ijerph19169922

Markov, K. K., Kudryavtsev, M. D., & Lmpetova, T. D. (2020). Healthy lifestyle cultivation and psychomotor progress model for elementary school. *Health Protecting Physical Education Technologies,* (5), 9-11. http://www.teoriya.ru/en/node/12231

Martinez-García, A., & Trescastro-López, E.M. (2016). Actividades de educación alimentaria y nutricional en escolares de 3º de primaria en el Colegio Público "La Serranica" de Aspe (Alicante): Experiencia piloto. *Revista Española de Nutrición Humana y Dietética 20*(2), 97-103. https://dx.doi.org/10.14306/renhyd.20.2.192

Mavilidi, M., Lubans, D., Morgan, P. J., Miller, A., Eather, N., Karayanidis, F., Lonsdale, C., Noetel, M., Shaw, K., & Riley, N. (2019). Integrating physical activity into the primary school curriculum: rationale and study protocol for the "Thinking while Moving in English" cluster randomized controlled trial. *BMC Public Health, 19*(379), 1-12. https://doi.org/10.1186/s12889-019-6635-2

McCullog, N., Boyle, S. E., Fothergill, M., & Defeyter, M. A. (2019). 'A really good balance': Thematic analysis of stakeholders' views on classroom- and games-based positive choices interventions for primary school children. *PLoS ONE, 14*(7), 1-15. https://doi.org/10.1371/journal.pone.0219503

Mcloughlin, G. M., Candal, P., Vazou, S., Lee, J. A., Dzewaltowski, D. A., Rosenkranz, R. R., Lanningham-Foster, L., Gentile, D. A., Liechty, L., Chen, S., & Welk, G. J. (2020). *Evaluating the implementation of the SWITCH® school wellness intervention and capacity-building process through multiple methods*. Springer Science and Business Media LLC. https://doi.org/10.1186/s12966-020-01070-y

Methley, A. M., Campbell, S., Chew-Graham, C., Mcnally, R., & Cheraghi-Sohi, S. (2014). *PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews*. Springer Science and Business Media LLC. https://doi.org/10.1186/s12913-014-0579-0

Moscoso-Loaiza, L. F., & Díaz-Heredia, L. P. (2018). Adopción de comportamientos saludables en la niñez: análisis del concepto. *Aquichan, 18*(2), 171-185. https://doi.org/10.5294/aqui.2018.18.2.5

Oberto, M. G., Mamondi, V., Ferrero, M., & Sánchez, R. J. (2020). Relato de una experiencia de promoción de la salud en escuelas: fomentando el lavado de manos. *Educación (Universidad De Costa Rica), 44*(2), 32. https://doi.org/10.15517/revedu.v44i2.37581

Organización Mundial de la Salud (OMS). (1999). *Healthy living: what is a healthy lifestyle?* Organización Mundial de la Salud. Recuperado 13.12.2022, de https://apps.who.int/iris/handle/10665/108180

Recibido: 27/04/2023 Aceptado: 22/12/2023

Organización Mundial de la Salud (OMS). (2020). *Directrices de la OMS sobre actividad física y hábitos sedentarios: de un vistazo*. Organización Mundial de la Salud. Recuperado 13.12.2022, de https://apps.who.int/iris/bitstream/handle/10665/337004/9789240014817-spa.pdf

Organización Mundial de la Salud (OMS). (2021). *La UNESCO y la OMS instan a los paises a que conviertan cada escuela en una escuela promotora de salud.* Organización Mundial de la Salud. Recuperado 13.12.2022, de https://www.who.int/es/news/item/22-06-2021-unesco-and-who-urge-countries-to-make-every-school-a-health-promoting-school

Organización Mundial de la Salud (OMS). (2022). *Noncommunicable diseases.* Organización Mundial de la Salud. Recuperado 15.12.2022, de https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & . . . McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration. Updated guidance and exemplars for reporting systematic reviews. *The British Medical Journal, 372*(160), 1-35. https://doi.org/10.1136/bmj.n160

Pulido-Gil, J. M., Sánchez-Oliva, D., López-Gajardo, M. Á, Ponce-Bordón, J. C., & García-Calvo, T. (2022). Effectiveness of a physically active learning program on indicators of physical activity, well-being and academic performance in students. *Cultura, Ciencia Y Deporte, 17*(52), 199-207. http://doi.org/10.1280 0/ccd.v17i52.1792

Real Decreto 157/2022, de 1 de marzo, por el que se establecen la ordenación y las enseñanzas mínimas de la Educación Primaria. *Boletín Oficial del Estado, 52*, 2 de marzo de 2022. https://www.boe.es/eli/es/rd/2022/03/01/157/con

Santos-Beneit, G., Bodega, P., de Miguel, M., Rodríguez, C., Carral, V., Orrit, X., Haro, D., Carvajal, I., de Cos-Gandoy, A., Peñalvo, J.L., Gómez-Pardo, E., Oliva, B., Ibañez, B., Fernández-Alvira, J.M., Fernández-Jiménez, R., & Fuster, V. (2019). Rationale and design of the SI! Program for health promotion in elementary students aged 6 to 11 years: A cluster randomized trial *American Heart Journal, 210,* 9-17. https://doi.org/10.1016/j.ahj.2018.12.011

Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making, 7*(16), 1-6. https://doi.org/10.1186/1472-6947-7-16

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A. (2015). Enhancing Cognitive and Social–Emotional Development Through a Simple-to-Administer Mindfulness-Based School Program for Elementary School Children: A Randomized Controlled Trial. *Developmental Psychology, 51*(1), 52-66. https://doi.org/10.1037/a0038454

Shum, A., Lai, E., Leung, W. G., Cheng, M., Wong, H. K., So, S., Law, Y. W., & Yip, P. (2019). A Digital Game and School-Based Intervention for Students in Hong Kong: Quasi-Experimental Design. *Journal of Medical Internet Research, 21*(4), 1-13. https://doi.org/10.2196/12003

Sitzer, D. L., & Stockwell, A. B. (2015). The art of wellness: A 14-week art therapy program for at-risk youth. *The Arts in Psychotherapy 45*, 69–81. http://dx.doi.org/10.1016/j.aip.2015.05.007

Stice, E., Shaw, H., y Marti, C. N. (2006). A meta-analytic review of obesity prevention programs for children and adolescents: The skinny on interventions that work. *Psychol Bull, 132*(5), 667–91. https://doi.org/10.1037/0033-2909.132.5.667

Suldo, S. M., Heaon, B. V., Bander, B., McCullough, M., Garofano, J., Roth, R. A., & Yin, S. (2015). Increasing Elementary School Students' Subjective Well-Being Through a Classwide Positive Psychology Intervention: Results of a Pilot Study. *Contemporary School Psychology, 19*, 300-311. https://doi.org/10.1007/s40688-015-0061-y

Vickery, C. E., & Dorjee, D. (2016). Mindfulness Training in Primary Schools Decreases Negative Affect and Increases Meta-Cognition in Children. *Frontiers in Psychology, 6,* 1-13. https://doi.org/10.3389/fpsyg.2015.02025

Watson, A., Timperio, A., Brown, H., & Hesketh, K. D. (2017). A primary school active break programme (ACTI-BREAK): study protocol for a pilot cluster randomised controlled trial. *Trials, 18*(433), 1-11. https://doi.org/10.1186/s13063-017-2163-5

Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D., & Dietz, W. H. (1997). Predicting obesity in young adulthood from childhood and parental obesity. *The New England Journal of Medicine, 337*(13), 869-873. https://doi.org/10.1056/NEJM199709253371301

Información de contacto: Olatz Arce-Larrory. Universidad de Deusto, Facultad de Educación y Deporte, Departamento de Educación. 48007 Bilbao, España. **E-mail: olatz.arce.larrory@deusto.es**TABLA IV. Objetivo y resultados generales de los estudios seleccionados *(Continuada)*

Pensamiento Crítico y Pedagogía Crítica: coincidencias y complementariedades

Critical Thinking and Critical Pedagogy: coincidences and complementarities

https://doi.org/10.4438/1988-592X-RE-2024-406-647

Ana Ayuste González
https://orcid.org/0000-0002-7196-3010
Institut de Recerca en Educació de la Universitat de Barcelona

Jaume Trilla Bernet
https://orcid.org/0000-0003-1254-0939
Institut de Recerca en Educació de la Universitat de Barcelona

Resumen

El artículo pone en relación dos importantes corrientes pedagógicas que han hecho de la crítica su seña de identidad: la del Pensamiento Crítico y la de la Pedagogía Crítica. El objetivo principal de este trabajo no será -como es más habitual en la reflexión académica- señalar las diferencias entre las corrientes estudiadas, sino localizar sus posibles coincidencias y complementariedades mediante la revisión de los textos más representativos de ambas corrientes. En la primera y la segunda parte del artículo presentamos cada una de las corrientes: tradiciones en las que se inscriben, autores de referencia, estilos intelectuales, aplicaciones prácticas… En la tercera parte, documentamos y comentamos tres aspectos en los que ambas corrientes pueden coincidir o complementarse. En primer lugar, el hecho de que tanto John Dewey como Paulo Freire -principales referentes de cada una de las corrientes- coincidían en la relevancia que otorgaban a un pensamiento crítico realmente activo: el norteamericano mediante el concepto de *experiencia*; el brasileño mediante el concepto de *praxis*. En segundo lugar, las aportaciones principales de las dos corrientes resultan complementarías, ya que cada una de ellas ha hecho énfasis en dimensiones

Recibido: 05/05/2023 Aceptado: 11/03/2024

distintas del pensamiento crítico. Una, conceptualizando rigurosamente este tipo de pensamiento y aportando métodos, técnicas, materiales y actividades para desarrollarlo. La otra corriente -la de la Pedagogía Crítica-, polarizando su reflexión en el sentido del pensamiento crítico: cuestionar la desigualdad y las relaciones de poder en la sociedad actual y entender la educación como una forma de transformación social. Por decirlo de otra forma: una corriente ha centrado su interés en el *qué* y la otra en el *para qué* del pensamiento crítico. Finalmente, las dos corrientes coinciden en una omisión: ambas han tenido muy poco presentes realizaciones pedagógicas europeas que han sido, *avant la lettre*, del todo coherentes con los planteamientos tanto del Pensamiento Crítico como de la Pedagogía Crítica.

Palabras clave: pensamiento crítico, pedagogía crítica, experiencia, praxis, pensamiento reflexivo.

Abstract

The article links two important pedagogical currents that have made criticism their hallmark: Critical Thinking and Critical Pedagogy. The main objective of this work is not to point out the differences between the two currents -as usually done in academic reflection-, but to identify their coincidences and complementarities through the review of the most representative texts of both currents. In the first and second parts of this article, we introduce each current: the traditions they belong to, reference authors, intellectual styles, practical applications… In the third part, we discuss three aspects in which both currents coincide or complement each other. Firstly, the fact that both John Dewey and Paulo Freire—the main referents of each current—agreed on the relevance of active critical thinking: the North American through the concept of experience; the Brazilian through the concept of praxis. Secondly, the main contributions of the two currents are complementary in that each one emphasizes different dimensions of critical thinking. Critical Thinking has rigorously conceptualized this type of thinking and provided methods, techniques, materials, and activities to develop it. Critical Pedagogy, has focused on the direction of critical thinking, questioning inequality and power relations in today's society and understanding education as a form of social transformation. In other words, while one current has focused on the "what" of critical thinking, the other one has focused on the "why". Finally, the two currents coincide in an omission: they have given very little consideration to European pedagogical achievements that have been, *avant la lettr*e, fully consistent with both Critical Thinking and Critical Pedagogy.

Keywords: critical thinking, critical pedagogy, critical pedagogy, experience, praxis, reflective thinking.

Introducción

Pensar críticamente, con autonomía e ideas propias, se ha convertido en una cualidad muy apreciada en distintos ámbitos (educativo, político, económico, cultural...) (Ennis, 2018; Facione, 2020). Nuestra sociedad anhela una ciudadanía crítica y comprometida con la mejora de la democracia y del bien común. La universidad, por su parte, afirma que quiere formar científicos e intelectuales críticos para que contribuyan a generar nuevos conocimientos. Y, por lo visto, hasta las empresas ahora desean que sus trabajadores piensen críticamente para resolver los problemas a los que deben enfrentarse, examinando todas las opciones que tienen a su alcance y así mejorar la productividad. Por todo ello, la promoción del pensamiento crítico es en la actualidad una de las finalidades importantes que se le plantean a la educación (Blair, 2021).

Sin embargo, la propia noción de pensamiento crítico y cómo han de formarse las personas capaces de pensar críticamente son, ya de entrada, cuestiones muy complejas y controvertidas. Y ahí entran en juego las dos corrientes objeto de este artículo: la del Pensamiento Crítico, arraigada en la tradición racionalista occidental, que pretende capacitar a los educandos para que piensen según unos estándares intelectuales muy exigentes; y la Pedagogía Crítica, con raíces en la tradición marxista, que se propone formar pensadores críticos para que asuman su condición de sujetos políticos y se comprometan con la transformación social. Ambas han querido conjugar lo crítico con lo educativo, pero constituyen, desde luego, propuestas bastante dispares: en sus tradiciones, autores de referencia, estilos intelectuales, métodos, aplicaciones prácticas… A primera vista, sólo tienen en común que las dos enarbolan lo "crítico" como signo de identidad. Son, además, muy escasos los nexos explícitos entre ellas, y tampoco suelen citarse mutuamente, sino es, en todo caso, para criticarse, como hace años documentaron muy bien Burbules y Berk (1999).

El texto al que acabamos de referirnos nos va a servir también para concretar el objeto específico de las páginas siguientes. Nuestro artículo guarda un cierto parentesco con el trabajo de Burbules y Berk, y es de justicia reconocer de entrada nuestra deuda con él. Se trata de un trabajo muy interesante, en el que sus autores mostraban un conocimiento profundo de las corrientes confrontadas. Pero reconocer esta deuda no presupone que nuestro artículo sea una réplica o comentario del otro.

Recibido: 05/05/2023 Aceptado: 11/03/2024

En ambos se pretende poner en relación la corriente del Pensamiento Crítico con la de la Pedagogía Crítica, pero hay una diferencia importante: Burbules y Beck se centraron sobre todo en las diferencias existentes entre las dos corrientes; el objetivo de nuestro artículo es, en cambio, hallar y argumentar sus posibles coincidencias. La intuición de la que partimos -y que querríamos comprobar- es que ambas corrientes, lejos de contradictorias o incompatibles, podrían ser, a pesar de sus diferencias, complementarias. Para ello, hemos realizado una lectura sistemática de los textos más representativos de ambas corrientes.

Los dos primeros apartados del artículo los dedicaremos respectivamente a presentar cada corriente. En el tercer apartado expondremos y documentaremos aquellas coincidencias y complementariedades. Un epílogo para reflexionar sobre lo hecho cerrará el artículo.

La corriente del Pensamiento Crítico

Los trabajos de John Dewey y el movimiento de educación progresista que se desarrolló en Norteamérica durante la primera mitad del siglo XX, contribuyeron a despertar el interés por la educación del pensamiento crítico (Ennis, 2011). Sin embargo, la denominación que utilizó Dewey en "How We Think"[1] -libro de cabecera para los autores de esta corriente- fue la de "pensamiento reflexivo", que Dewey valoraba como la mejor manera de pensar y que describía como ese "darle vueltas a un tema en la cabeza y tomárselo en serio con todas sus consecuencias" (Dewey, 1989, p. 21). El pensamiento reflexivo, según el filósofo y pedagogo, impulsa a la investigación y consiste en "el examen activo, persistente y cuidadoso de toda creencia o supuesta forma de conocimiento a la luz de los fundamentos que la sostienen y las conclusiones a las que tiende" (1989, p. 25).

Dewey es una referencia constante para la corriente del Pensamiento Crítico, pero según uno de sus autores más cualificados, Robert H. Ennis, fueron diversos hechos acontecidos ya en la segunda mitad del siglo XX -la crisis del Sputnik (1957) o los movimientos de protesta contrarios

[1] El título completo del libro es *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*, y su primera edición fue en 1910. *How we think*. Boston: D. C. Heath and Company Publishers. En lo que sigue utilizaremos la versión española de 1989.

Recibido: 05/05/2023 Aceptado: 11/03/2024

a la guerra de Vietnam de los años 60, entre otros- lo que renovó el interés por el pensamiento crítico, provocando que muchas instituciones educativas norteamericanas se comprometieran con su investigación y enseñanza (Ennis, 2011, 2018). Este creciente interés y la falta de una visión clara sobre qué habilidades había que enseñar y cómo enseñarlas y evaluarlas llevó, a finales de los 80, a que la American Philosophical Society pidiera a Peter Facione, un filósofo de la educación, que investigara sobre el tema (Fisher, 2021). Facione reunió a un panel de 46 expertos de Estados Unidos y Canadá, muchos de ellos implicados en organizaciones fundadas específicamente para promover el pensamiento crítico. Durante dos años, estos expertos, entre los que se encontraban Robert Ennis; Richard Paul, Robert Swartz o Harvey Siegel, se dedicaron a debatir, mediante el método Delphi, sus puntos de vista con la finalidad de acordar un posicionamiento común. El resultado de este proceso de discusión culminó en una declaración conjunta titulada *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*, en la que la definición de consenso alcanzada ponía de relieve que el pensamiento crítico requiere tanto de unas habilidades cognitivas -interpretación, análisis, evaluación, inferencia, autorregulación-, como de una serie de disposiciones o hábitos mentales: a) Curiosidad dirigida a una amplia variedad de temas y asuntos. b) Preocupación por estar y mantenerse bien informado. c) Estado de alerta para detectar oportunidades de utilización del pensamiento crítico. d) Confianza en los procesos de investigación razonados. e) Autoconfianza en las propias habilidades para razonar. f) Mente abierta respecto a visiones divergentes. g) Flexibilidad al considerar opiniones de otros y alternativas. h) Comprensión de las opiniones de otras personas. i) Imparcialidad en la valoración del razonamiento. j) Honestidad al enfrentar las propias predisposiciones, prejuicios, estereotipos. k) Prudencia al postergar, realizar o alterar juicios. l) Voluntad para reconsiderar y revisar visiones (Facione, 1990).

Desde esta perspectiva, la educación de buenos pensadores críticos consiste en la combinación de las habilidades cognitivas mencionadas con las disposiciones que fomentan la introspección y que son, a su vez, la base de una sociedad racional y democrática (Facione, 1990). Por esta razón, el egocentrismo es visto por estos autores como una limitación del pensamiento crítico. Las personas, como sugieren Paul y Elder: "no solemos considerar los derechos y necesidades de los demás, ni solemos apreciar el punto de vista de los otros o las limitaciones de

Recibido: 05/05/2023 Aceptado: 11/03/2024

nuestro punto de vista. Nos damos cuenta de nuestro pensar egocentrista cuando nos adiestramos para ello" (2003, p. 9). Por eso es preciso distinguir entre pensamiento crítico "débil" y "fuerte" (Paul, 1981). Las personas que se guían por un sentido "débil" emplean las habilidades críticas para reforzar la propia posición; en cambio cuando se inclinan por un sentido "fuerte", las emplean para dilucidar sobre sus propios procesos cognitivos y afectivos, precisamente en aquellos temas en los que se es más proclive a sesgos egocéntricos o sociocéntricos. Enseñar a pensar críticamente en el sentido débil conduce a que el estudiante se vuelva más sofista, más hábil para racionalizar sus propios sesgos y para mantener sus prejuicios o hábitos irracionales. Pensar críticamente en un sentido fuerte le da un carácter autorreflexivo y dialógico. Implica el intercambio entre diferentes puntos de vista, marcos de referencia, interpretaciones, etc.

Esta distinción (pensamiento crítico *fuerte/débil*) que proponía Paul hace más de cuarenta años, resulta especialmente oportuna ante fenómenos actuales como la eclosión de las redes sociales, las *fake news* y, en general, el desarrollo de las tecnologías mediáticas que ponen a disposición de amplísimos sectores de la ciudadanía la posibilidad de ejercer la crítica de forma pública (Comas *et al.*, 2021; Cañas *et al.*, 2022). Una crítica que antes estaba reservada a los sectores sociales, muy minoritarios (políticos, periodistas, intelectuales mediáticos…), que tenían acceso a los medios de comunicación convencionales. Ahora, las nuevas tecnologías y plataformas posibilitan que sectores muchísimo más amplios compartan públicamente su pensamiento crítico. Esto es, sin duda, un logro democrático. Pero, a la vez, obliga a los usuarios a un ejercicio constante de discriminación, pues ahí puede haber -hay- crítica de todo tipo: crítica basura, descalificaciones insultantes, opiniones indocumentadas, *fake news*… (Innerarity, 2022). Todo lo cual ya existía antes, pero los medios convencionales podían hacer una cierta criba, la cual suele ser mucho más laxa en las plataformas actuales. Esto obliga a que sea el receptor quien tenga que discriminar entre los pensamientos críticos que se le ofrecen. Una tarea educativa ya urgente es formar a la ciudadanía para esta función selectiva entre la multitud de "pensamientos críticos" a su alcance. Las categorías propuestas por Paul son ya un primer paso para disponer de criterios útiles para esta tarea de discriminación.

Así pues, para los autores de la corriente del Pensamiento Crítico, formar personas críticas requiere desarrollar una serie de habilidades intelectuales, adquirir conciencia y control sobre los procesos de

Recibido: 05/05/2023 Aceptado: 11/03/2024

pensamiento y aprendizaje para mejorar dichas habilidades, y formar aquellas disposiciones o hábitos mentales que producen reflexiones útiles. De ahí que la corriente del Pensamiento Crítico se identifique, según Facione (2020), con el modelo de educación liberal que consiste en aprender a aprender y aprender a pensar por uno mismo, tanto de forma individual como en colaboración con otros, para lo que algunos miembros de la misma han elaborado diferentes propuestas y guías educativas (Lipman, 1998; Ennis, 2005; Paul y Elder, 2007; Facione, 2020). El pensamiento crítico se convierte en un pilar fundamental de las sociedades racionales y democráticas puesto que supone avanzar desde una posición centrada en la toma de decisiones de carácter personal, hacia otra que tenga en cuenta la preocupación por el bien común y la justicia social (Facione, 2020; Paul y Eder, 2007; Ennis 2018).

La corriente de la Pedagogía Crítica

Así como la corriente anterior convirtió a John Dewey en su referente por antonomasia, los autores de la Pedagogía Crítica realizaron una operación semejante con Paulo Freire, aunque con dos diferencias significativas: una propiciada por la cronología y la otra de carácter contextual. El tiempo de Freire y el de los autores de la Pedagogía Crítica coincidió en parte, lo cual posibilitó que algunos de estos autores pudieran mantener un contacto directo, continuado e incluso personal con quien habían elegido como referente (Giroux, 1990; Apple 1996; Kincheloe, 2004; Darder, 2017); eso mismo, por obvias razones cronológicas, no pudo ocurrir en el caso de los autores del Pensamiento Crítico con John Dewey.[2] Por otro lado, si bien es cierto que ambas corrientes se han desarrollado fundamentalmente en Norteamérica -después retomaremos esta cuestión-, el gran referente que eligió la Pedagogía Crítica hizo sus primeras y más relevantes aportaciones en y para un contexto social (el denominado tercer mundo) muy diferente al propio de los principales autores de la Pedagogía Crítica (primerísimo mundo). Eso mismo ha sido incluso tematizado por algunos de estos autores (McLaren y Kincheloe, 2008).

[2] Es preciso anotar que, aunque Freire ha sido siempre la referencia principal de la Pedagogía Critica, esta corriente también ha valorado positivamente las aportaciones de Dewey (Apple y Beane, 2000; Kincheloe, McLaren y Steinberg, 2003; Giroux, 2020; McLaren, 2022).

Recibido: 05/05/2023 Aceptado: 11/03/2024

Giroux (1990), Apple (1996), Macedo (2000), McLaren y Kincheloe (2008), Darder (2017), ..., encontraron en la pedagogía del autor brasileño una forma de combinar el *lenguaje de la crítica* con el *lenguaje de la posibilidad* (Giroux, 1990). Combinación que les permitió describir la educación como una institución de dos caras, con finalidades y resultados diferentes. Por un lado, la educación regida por un programa estandarizado que sirve a los intereses del poder dominante y de aquellos estudiantes más íntimamente ligados con los estándares sociales y culturales que se asocian a dicho poder. Y, por otro, una educación democrática sensible a las diferencias que persigue la inclusión y la promoción social y cultural, especialmente de los estudiantes que provienen de grupos desfavorecidos. El programa de la Pedagogía Crítica se erigió así sobre la crítica a las instituciones educativas al servicio del sistema capitalista como las teorías de la reproducción denunciaban (Althusser, Bourdieu y Passeron, Baudelot y Establet, Bowles y Gintis, …), pero también proyectando nuevas maneras de entender la acción educativa para cambiar las relaciones de poder y las situaciones de injusticia social derivadas de dicho sistema.

En un contexto marcado por las luchas de liberación colonial y nacional, las protestas estudiantiles y la aparición de una pluralidad de movimientos sociales (feminismo, derechos civiles, pacifismo, ecologismo…), la publicación de *Pedagogía del Oprimido* de Paulo Freire en 1968 y su posterior traducción al inglés en 1970, fue fuente de inspiración para aquellas comunidades y educadores en Estados Unidos que luchaban por un mundo libre de injusticias y dominación (Kincheloe, 2008, Darder, 2017). La praxis político-pedagógica de la pedagogía freiriana hacía posible superar los dualismos entre sujeto y objeto, persona y mundo, conciencia y realidad, y teoría y práctica; y, asimismo, concebir la educación como un proceso de concientización mediante un diálogo horizontal en el que los sujetos dialogan con el mundo y entre sí para problematizar su realidad y transformarla. (Trilla, 1989; Ayuste y Trilla, 2005).

Desde esta perspectiva e influidos por Gramsci y la Escuela de Frankfurt, la razón de ser de la educación es formar una ciudadanía crítica capaz de construir una sociedad igualitaria y realmente democrática. Para ello, los trabajos de la Pedagogía Crítica han consistido en examinar cómo el modo de producción capitalista -y también el patriarcado y el colonialismo- inciden en las políticas y prácticas educativas; y, al mismo

Recibido: 05/05/2023 Aceptado: 11/03/2024

tiempo, profundizar en los principios de la educación democrática y en el papel de los educadores como intelectuales que desarrollan pedagogías contrahegemónicas (Apple y Beane, 2000; Giroux, 1990). Las pedagogías críticas, según Giroux (1990), "no sólo potencian a los estudiantes proporcionándoles el conocimiento y las habilidades sociales que necesitarán para actuar en el conjunto de la sociedad con sentido crítico, sino que, además, los educan para la acción transformadora" (p. 35). Y es este énfasis en educar para la acción transformadora uno de los aspectos que más distancia la Pedagogía Crítica de la corriente del Pensamiento Crítico.[3]

Otra aportación de la Pedagogía Crítica, según McLaren (2022), ha sido la de cuestionar la concepción del conocimiento como algo "neutral" u "objetivo" y tratar de comprender la naturaleza política de la educación. De ahí que una de las tareas principales de esta corriente resida en intentar conocer, desde abajo o, en otras palabras, desde la perspectiva de los propios oprimidos, los mecanismos de opresión impuestos por el orden establecido. Sin embargo, con el paso del tiempo, como nos advierte el mismo autor, la Pedagogía Crítica se ha vuelto más ecléctica y se ha centrado menos en la crítica al capitalismo y a las desigualdades que produce. El alcance de la pedagogía crítica se extiende ahora a la educación multicultural, la educación bilingüe, el aprendizaje de idiomas y la alfabetización crítica (incluida la mediática). Y, desde entonces, sus principales exponentes han cruzado a la pedagogía crítica con diferentes tradiciones transdisciplinares, incluyendo incursiones teóricas en los trabajos de autores como Rorty, Lacan, Derrida y Foucault. Un hecho que, según Wheeler-Bell (2019), ha supuesto para la Pedagogía Crítica una crisis de identidad.

[3] En el trabajo de Burbules y Berk (1999) citado al principio del artículo, estos autores dan buena cuenta de las diferencias entre las dos perspectivas. Sin ánimo de ser exhaustivos, las diferencias más sustanciales que señalan estos autores podrían sintetizarse en las siguientes. Para la corriente del Pensamiento Crítico, como veíamos, la finalidad principal de la educación sería desarrollar aquellas habilidades y disposiciones que permitan a la persona pensar por sí misma según unos estándares intelectuales exigentes. Esto, para la Pedagogía Crítica, sería insuficiente dado que el objetivo último de la educación no radica sólo en comprender la realidad sino en transformarla. Para ello, la alfabetización política es un requisito básico. Algo que los del pensamiento crítico rechazan porque, según ellos, comporta un alto riesgo de adoctrinamiento. Los contenidos propuestos por una y otra corriente son también diferentes. Para la Pedagogía Crítica han de estar estrechamente relacionados con la experiencia y el contexto de los estudiantes. Para la otra perspectiva, justamente la familiaridad con el contexto y los problemas a abordar pueden acentuar los sesgos culturales y limitar la imparcialidad.

Recibido: 05/05/2023 Aceptado: 11/03/2024

En una era marcada por el darwinismo económico neoliberal, la colonización mercantilista de la vida cotidiana y una nueva clase de fascismo, en palabras de Giroux (2022), es urgente para estos autores poner de nuevo la educación en el centro de la política. La Pedagogía Crítica como forma de política cultural ofrece la "promesa de un espacio protegido dentro del cual es posible pensar a contracorriente" (Giroux, 2022, p. 201). Como parte del lenguaje de la esperanza, la Pedagogía Crítica sigue aspirando a convertirse en un marco de referencia en el que teoría y práctica puedan confluir para producir formas alternativas de enseñar y aprender, conectando la educación con los asuntos que afectan a la ciudadanía y la mejora de la democracia (McLaren y Kincheloe, 2008; Giroux, 2020).

Coincidencias y complementariedades

Resumido lo esencial de ambas corrientes, podemos llegar ya al objeto más específico del artículo: elucidar las coincidencias y complementariedades existentes o posibles entre Pensamiento Crítico y Pedagogía Crítica. Lo haremos en tres apartados. En el primero constataremos una coincidencia muy especial pues afecta a los referentes clave de las dos corrientes. En un segundo momento veremos que cada corriente ha enfatizado aspectos diferentes del pensamiento crítico, lo cual -en una lectura positiva- convertiría en complementarias a estas corrientes. Finalmente, pondremos en evidencia otra coincidencia; aunque en este caso se tratará de una omisión compartida.

Un punto de encuentro: Dewey y Freire, experiencia y praxis

John Dewey y Paulo Freire, según veíamos, son las referencias pedagógicas ineludibles, respectivamente, para cada una de las dos corrientes: El libro *How we think* (1910) del filósofo y pedagogo norteamericano es de cita casi obligada en los trabajos de la corriente del Pensamiento Crítico. Por su parte, la corriente de la Pedagogía Crítica no puede dejar de referirse a Paulo Freire.

No se trata de emparejar ahora las ideas y propuestas educativas de ambos autores. Los contextos históricos, sociales y políticos en los que (y

Recibido: 05/05/2023 Aceptado: 11/03/2024

para los que) trabajaron fueron, desde luego, muy diferentes; y por eso las respuestas que aportaron tampoco podían ser las mismas (Corbett y Guilherme, 2021). Ello no quita que entre ambos no puedan hallarse puntos de encuentro.

De entrada, ambos autores coinciden en la relevancia que la pedagogía contemporánea ya les ha reconocido, aunque para épocas parcialmente diferentes: Dewey fue, sin duda, uno de los más grandes de la pedagogía de la primera mitad del siglo pasado; mientras que Freire fue lo mismo, pero durante su segunda mitad. Y ambos -en esto también coinciden- siguen siendo ahora referentes importantes, tanto para la reflexión pedagógica como para la práctica educativa. Pero más allá de valoraciones de este tipo, hay un aspecto en el que cabe encontrar una analogía entre los dos autores que resulta muy significativa para nuestro tema. Se trata de la importancia que ambos otorgaron al aspecto fundamentalmente *activo* del pensamiento crítico; aunque cada uno de ellos manejando terminologías diferentes: *experiencia*, en el caso del norteamericano; *praxis*, en el caso del brasileño.

Es bien conocido el papel de la idea de *experiencia* en la concepción educativa de Dewey. Por de pronto, en la definición de educación formulada en su obra capital, *Democracia y educación*, "experiencia" es el elemento más nuclear: [educación] "es aquella reconstrucción o reorganización de la experiencia que da sentido a la experiencia y que aumenta la capacidad para dirigir el curso de la experiencia subsiguiente" (Dewey, 2004, p.74). Es decir, nos educamos por medio de la experiencia y nunca de otra forma. De hecho, con este concepto de educación como reconstrucción de la experiencia, lo que hizo Dewey fue complejizar y dar profundidad al famoso lema del *learnig by doing*, pero sin desmentirlo en absoluto: aprendemos *haciendo*, pero no haciendo cualquier cosa. Las experiencias que verdaderamente nos educan son aquellas que, partiendo de nuestras experiencias previas, consiguen que ahora las comprendamos mejor y que, a la vez, nos sirvan también para orientar las experiencias siguientes.

Así pues, el concepto de experiencia, tal y como lo entiende Dewey, es lo que permite establecer una relación mutuamente enriquecedora entre la teoría y la práctica, entre la acción y la reflexión. ¿Y no es eso algo muy parecido a lo que también pretendía Freire mediante el concepto de *praxis*? Como explicábamos en otro lugar (Trilla, 1993, 151-175) y también otros resaltaron (Gadotti, 1996), *praxis* es un concepto fundamental en el pedagogo brasileño: "Separada de la práctica, la teoría

es puro verbalismo inoperante; desvinculada de la teoría, la práctica es activismo ciego. Es por esto mismo que no hay praxis auténtica fuera de la unidad dialéctica acción-reflexión, práctica-teoría» (Freire, 1984, p. 30). Y, en otro lugar, introduciendo los conceptos de *contexto teórico* y *contexto práctico*, Freire escribía:

> Es necesario que la praxis, a través de la cual la conciencia se transforma, no sea pura acción, sino acción y reflexión. (…) La praxis teórica no es otra cosa que aquello que hacemos desde el contexto teórico, al tomar distancia con respecto a la praxis realizada o que está realizándose en el contexto concreto. Por ello es que la praxis teórica sólo es auténtica en la medida en que no se rompa el movimiento dialéctico entre ella y la subsiguiente praxis a ser realizada en el contexto concreto. De ahí que ambas formas de praxis sean momentos indicotomizables de un mismo proceso por el cual conocemos en términos críticos. (Freire, 1975, pp. 12-13).

Así pues, según Freire, el concepto de *concientización* -fundamental en su concepción pedagógica, como es bien sabido- implica algo más que una simple toma de conciencia de la propia realidad. Es, en todo caso, una toma de conciencia que pretende ser objetiva, crítica y comprometida. Una toma de conciencia que solamente es posible adquirir mediante la praxis, y nunca solamente a través de una pura actividad intelectual desligada de la acción concreta (Freire, 1985).

Las experiencias que, según Dewey, verdaderamente nos educan serían aquellas -puestas ahora en términos freirianos- que nos hacen más críticamente conscientes de la realidad en la que estamos y actuamos, y nos sirven para dirigir nuestras praxis subsiguientes. En épocas y contextos diferentes y con terminología igualmente distinta, los referentes clave de cada una de las corrientes coincidían en el carácter esencialmente *activo* (experiencial/práxico) de la educación y el pensamiento.

Énfasis diferentes

Decíamos que la corriente del Pensamiento Crítico y la de la Pedagogía Crítica podrían ser también complementarias, pues cada una de ellas ha hecho énfasis en aspectos diferentes. Por lo que se refiere a la corriente del Pensamiento Crítico, destaca su esfuerzo por conceptualizar

Recibido: 05/05/2023 Aceptado: 11/03/2024

rigurosamente este tipo de pensamiento. El hecho de que los estudios sobre el pensamiento crítico se hayan realizado principalmente desde la filosofía analítica, ha motivado que algunos autores no hayan podido evitar preguntarse también sobre el pensamiento autocrítico o sobre cómo el pensamiento crítico es capaz de pensarse a sí mismo. Y, en un nivel ya más estrictamente pedagógico, la corriente del Pensamiento Crítico también ha aportado métodos, técnicas, actividades y materiales, algunos muy concretos, para desarrollar este tipo de pensamiento en las personas: cómo aprender a pensar críticamente (Lipman, 1998; Ennis, 2005; Paul y Elder, 2007; Facione 2020). En este sentido, López (2012) ha distinguido cuatro modelos de enseñanza del pensamiento crítico: el modelo *de evaluación procesual*, el modelo de *pensamiento dialógico*; el modelo de *comunidad de investigación* y el modelo de *controversia*. Del mismo modo que también se han creado herramientas para evaluar el pensamiento crítico (Saiz y Rivas, 2008).

En cambio, parece que el movimiento de la Pedagogía Critica ha centrado su interés sobre todo en los contenidos valorativos y proyectivos del pensamiento crítico. Por decirlo así: pensar críticamente nuestra realidad social; y pensar en la sociedad que debería construirse. Si las disciplinas académicas más asociadas a la corriente del Pensamiento Crítico eran la Psicología y la Filosofía (sobre todo la analítica y el pragmatismo), la Pedagogía Crítica se encuentra mucho más vinculada a la Sociología; y también a la Filosofía, pero en este caso a la desarrollada por autores marxistas o neomarxistas como los de la Escuela de Frankfurt. Aunque más adelante, la Pedagogía Crítica también empezó a nutrirse del postestructuralismo con vistas a alcanzar un concepto de agencia humana que abarcara el género y la cultura y no se limitara a la clase social (Wheeler-Bell, 2019; Ichikawa, 2020).

Retomando nuestra hipótesis sobre la complementariedad de ambas corrientes, a veces parece que los de la Pedagogía Crítica flojeen en el *cómo*, mientras que los del Pensamiento Crítico a veces titubeen un poco en el *para qué*. Es verdad que los primeros siempre tienen a mano el recurso de Freire, pues él sí que aportó realizaciones, métodos y estrategias. Pero obviando las realizaciones del pedagogo brasileño y unas pocas experiencias más a las que también se han referido a veces (Apple y Beane, 2000), leyendo a los de la Pedagogía Crítica a menudo se encuentran a faltar los ejemplos, los proyectos, los métodos, las técnicas, las actividades concretas para llevar a cabo sus magníficos planteamientos

sobre la transformación social por medio de la educación. La Pedagogía Crítica, hasta ahora, ha consistido más en discursos que en prácticas. Con los del Pensamiento Crítico ocurre, en cambio, un poco todo lo contrario: ellos sí que, como decíamos, ofrecen técnicas, materiales y actividades para desarrollar el pensamiento crítico, pero queda mucho menos claro el *para qué*. Es como si dieran por supuesto que el mundo ya va a ser mejor por el hecho de que haya más personas que piensen críticamente. Algunos autores encuadrables en esta corriente, como veíamos antes, al final acaban casi por identificar pensamiento crítico con pensamiento científico. Pensar críticamente sería pensar la realidad (describirla, explicarla, comprenderla…) lo más rigurosa y objetivamente posible; casi podríamos decir que, para ellos, pensar críticamente es *pensar bien*; la pregunta pendiente sería: pensar bien, ¿pero para criticar qué? Estas diferencias, en cierto modo, convierten en complementarias a las dos corrientes: una, porque aporta contenidos para entender mejor el qué y el cómo del pensamiento crítico; la otra porque responde más claramente al para qué del mismo.

En relación a todo esto, no deja de ser curioso (o paradójico) que actualmente a lo que más suele asociarse el pensamiento crítico es a la formación para la emprendeduría. No hay más que poner en Google "pensamiento crítico" y "emprendeduría" (o cualquiera de los derivados de esta palabra tan de moda) para ver como aparecen centenares de miles de páginas web, así como libros enteros[4] y artículos dedicados a magnificar la necesidad de que los emprendedores piensen críticamente. Y eso no sólo ocurre en internet y en el discurso empresarial, sino también en el propio discurso pedagógico y político. Una entidad supranacional tan importante como el Consejo de la Unión Europea en sus recomendaciones sobre las "Competencias clave para el aprendizaje permanente" (2018), menciona expresa y literalmente al pensamiento crítico sólo en tres de las ocho competencias clave propuestas;[5] una de las tres es, por supuesto, la "competencia emprendedora".[6] Es posible que estas asociaciones tan claras y directas entre el pensamiento crítico

[4] *Kant en Silicon Valley: Pensamiento crítico para emprendedores* de Rais Busom, es el título de uno de estos libros que circula por internet.

[5] Las ocho competencias propuestas son: "competencia en lectoescritura; competencia multilingüe; competencia matemática y competencia en ciencia, tecnología e ingeniería; competencia digital; competencia personal, social y de aprender a aprender; competencia ciudadana; competencia emprendedora; competencia en conciencia y expresión culturales".

[6] Las otras dos son: la competencia en lectoescritura y la competencia digital.

Recibido: 05/05/2023 Aceptado: 11/03/2024

y el discurso tecnocrático-empresarial escandalizara a los autores más radicales de la Pedagogía Crítica, pero no necesariamente a los del Pensamiento Crítico. Aquellos tendrían clara la respuesta a la pregunta anterior: pensamiento crítico, pero para criticar la desigualdad inherente a la sociedad de clases, para denunciar las relaciones de poder en el capitalismo, para hallar y poner en práctica alternativas transformadoras a la educación bancaria que perpetua la opresión de los oprimidos… Desde luego, hay diferencias entre la corriente del Pensamiento Crítico y la de la Pedagogía Crítica; sobre todo, entre los autores más políticamente radicales de la segunda y los más políticamente neutros de la primera.

Una omisión compartida. Pedagogías homologables *avant la lettre*

Pero volvamos a las coincidencias entre ambas corrientes. Tanto la una como la otra son de origen norteamericano, con la incorporación en la Pedagogía Crítica de un singular referente latinoamericano (Freire). En cualquier caso, las dos corrientes proceden del otro lado del Atlántico. Y eso se nota: es un sesgo en el que han caído, a pesar de la crítica que ellas mismas han realizado a cualquier clase de "centrismo" que enturbie la necesaria objetividad de un pensamiento crítico *fuerte*. Este centrismo americano quizá explique el hecho de que hayan sido muy escasas las referencias europeas, estrictamente pedagógicas, presentes en ambas corrientes.[7] Aludiremos por ello a algunas experiencias y movimientos educativos europeos que creemos plenamente homologables a los planteamientos tanto del Pensamiento Crítico como de la Pedagogía Crítica. Para no alargarnos, nos vamos a referir sólo a dos de ellas que gozan, además, de algunas características compartidas remarcables. La primera es que han tenido una proyección internacional notable, lo cual convierte aún en más extraño que, con algunas excepciones, los autores más significados de las dos corrientes americanas no las hayan considerado como antecesores o, al menos, como compañeros de viaje

[7] Es verdad que a la Pedagogía Crítica se la ha relacionado con Gramsci, la Escuela de Frankfurt, pero ello ha sido a nivel filosófico y no estrictamente pedagógico. Y también es verdad que hay autores que han ubicado a las teorías educativas de la reproducción europeas en la Pedagogía Crítica. Pero justamente lo que la Pedagogía Crítica ha intentado superar son este tipo de planteamientos reproduccionistas cerrados y circulares que niegan cualquier posibilidad de que la educación pueda ser un factor determinante de transformación social.

Recibido: 05/05/2023 Aceptado: 11/03/2024

de sus propios planteamientos. La otra característica remarcable es que las dos pedagogías europeas proceden de experiencias prácticas, reales; no son sólo discursos.

Por orden de antigüedad, hay que mencionar en primer lugar a Francisco Ferrer y Guardia, su Escuela Moderna y el movimiento de la Pedagogía Racionalista que desencadenó. ¿Puede caber alguna duda de que el desarrollo del pensamiento crítico más genuino era uno de los objetivos clave de aquella escuela barcelonesa de principios del siglo XX? ¿Y que la transformación social mediante la educación era también lo que perseguían Ferrer y sus seguidores?:

> Demostrar a los niños que mientras un hombre depende de otro hombre se cometerán abusos y habrá tiranía, y esclavitud, estudiar las causas que mantienen la ignorancia popular, conocer el origen de todas las prácticas rutinarias que dan vida al actual régimen insolidario, fijar la reflexión de los alumnos sobre cuanto a la vista se nos presenta, tal ha de ser el programa de las escuelas racionalistas. (…) La enseñanza racional y científica ha de persuadir a los futuros hombres y mujeres que no han de esperar nada de ningún ser privilegiado (ficticio o real); y que pueden esperar todo de lo racional de sí mismos y de la solidaridad libremente organizada y aceptada. (Ferrer y Guardia, 1976, p. 103-14)

Bastante de lo reivindicado por la actual Pedagogía Crítica lo podemos hallar ya en aquella pedagogía de hace más de cien años: la coeducación de sexos (un verdadero escándalo entonces); la lucha por la igualdad social y contra la exclusión, que Ferrer planteaba mediante la coeducación de las clases sociales; o un feminismo realmente radical para la época:

> La mujer no debe estar recluida en el hogar. El radio de su acción ha de dilatarse fuera de las paredes de las casas: debería ese radio concluir donde llega y termina la sociedad. Más para que la mujer ejerza su acción benéfica, no se han de convertir en poco menos que en cero los conocimientos que le son permitidos: debieran ser en cantidad y calidad los mismos que el hombre se proporciona. (Ferrer y Guardia, 1976, p.52)

La Escuela Moderna fue Pedagogía Crítica y Pensamiento Crítico de pura cepa. Como también lo fue aquella escuela perdida en un pueblecito de la Toscana italiana: Barbiana. Una aldea a la que fue desterrado -por rebelde, contestario y crítico- el cura Don Lorenzo Milani. *Carta a una*

Recibido: 05/05/2023 Aceptado: 11/03/2024

maestra, el pequeño libro escrito por los alumnos de la escuela, guiados por el cura, es un ejemplo diáfano de Pensamiento Crítico y de Pedagogía Crítica al unísono. Por de pronto, el librito es una muestra anticipada, pero muy fidedigna, de las *teorías de la reproducción*: la pedagogía excluyente que encarna la *maestra* a la que los jóvenes de Barbiana dirigen su *carta*; y el sistema escolar clasista (descrito incluso con estadísticas) del que habían sido expulsados aquellos niños y adolescentes. Pero el librito, contándonos como se trabajaba intensamente en la escuela de Barbiana, muestra también que la función reproductora del sistema escolar no tiene por qué ser inexorable; demuestra que hay alternativas pedagógicas factibles y que la educación puede ser un verdadero factor de transformación social. El propio libro de aquellos alumnos es pensamiento crítico en acción; un pensamiento crítico ejercitado, entre otros procedimientos, mediante la cotidiana lectura comentada de la prensa. Afirmaba un alumno de Barbiana:

> (…) sabía bien la historia en que yo vivo. Es decir, el periódico que leíamos en Barbiana todos los días en voz alta, de punta a cabo. (…) Nada del periódico sirve para vuestros exámenes. Es la prueba de que hay poco en vuestra escuela que sirve para la vida. Precisamente por eso hay que leerlo. (Alumnos de la escuela de Barbiana, 2017, p. 20)

También, en *Carta a una maestra* se afirma que "sólo la lengua nos hace iguales. Igual es quien sabe expresarse y entiende la expresión ajena. Que sea rico o pobre importa menos. Basta con que hable." (Alumnos de la escuela de Barbiana, 2017, p. 7). Milani murió un año antes de que viera la luz *Pedagogía del oprimido* (1968, publicado en español por primera vez en 1970); o sea, que el cura italiano no pudo conocer la pedagogía de Freire,[8] pero no hay duda de que Barbiana es un caso clarísimo de concientización *avant la lettre*. Milani y Freire no llegaron a conocerse personalmente, pero seguro que hubieran congeniado.

Por supuesto que, además de Barbiana y Ferrer y Guardia, habría otros muchos ejemplos posibles de pedagogías europeas homologables -anteriores o coetáneas- a muchas de las propuestas tanto de la

[8] Como explica el profesor José Luís Corzo (2021), Freire sí que posteriormente llegó a conocer la obra pedagógica de Milani. Freire escribió el "postfazione" al libro coordinado por Corzo y Gesualdi -uno de los alumnos de Barbiana-: *Don Milani nella scrittura collettiva* (1992).

Pedagogía Crítica como del Pensamiento Crítico. Tan indiscutible como las experiencias comentadas de Ferrer y Barbiana, podría ser la pedagogía popular que el francés Celestin Freinet fue pergeñando a partir de su propia práctica. Igualmente, nuestro eurocentrismo -paralelo al americanismo de las corrientes que nos han ocupado- nos ha impedido conocer otros planteamientos pedagógicos ajenos a estos dos centrismos, pero que seguramente coincidirían en mucho con las corrientes críticas estudiadas.

Haber puesto de relieve que las dos corrientes estudiadas omiten referencias a aportaciones y experiencias pedagógicas europeas relevantes plenamente homologables a los planteamientos de ambas, nos pone en la pista de algunos de los límites de nuestro trabajo y, por tanto, también de posibles prolongaciones del mismo. Cabría, por supuesto, profundizar en el estudio de cada una de las pedagogías a las que acabamos de referirnos (Escuela Moderna, Barbiana, Freinet), en tanto que precedentes, plenamente acreditados, de cómo pueden conjugarse, en la práctica, las aportaciones de las dos corrientes objeto del artículo. En los párrafos anteriores -por mor de la extensión debida del artículo- nos hemos limitado a justificar y fundamentar mínimamente que los ejemplos citados constituyen experiencias reales del todo encuadrables -aunque, *avant la lettre*- en los planteamientos principales tanto de la corriente de la Pedagogía Crítica como de la del Pensamiento Crítico. Tendría pleno sentido y justificación dedicar otros artículos (o incluso trabajos de mayor envergadura) a estudiar monográficamente cualquiera de aquellas experiencias, autores y movimientos, desde los discursos aportados por ambas corrientes. Son, desde luego, experiencias ya muy estudiadas, pero hacerlo ahora desde el bagaje de ideas y aparatos conceptuales del Pensamiento Crítico y de la Pedagogía Crítica redundaría en beneficio tanto de aquellas experiencias como de tales corrientes. Por un lado, pondría de relieve la actualidad de experiencias que, aunque fraguadas hace mucho tiempo, pueden seguir realizando aportaciones bien prácticas para el desarrollo del pensamiento crítico. Por otro lado, las propias corrientes objeto de nuestro artículo saldrían reforzadas con estudios como los propuestos. Hemos visto que no siempre ofrecían ejemplos concretos y prácticos de sus aportaciones teóricas: las experiencias y movimientos mencionados se los aportarían en cantidad.

Y, como prolongaciones posibles de este trabajo, lo mismo que decimos respecto de los tres ejemplos propuestos, cabría decirlo de otras

Recibido: 05/05/2023 Aceptado: 11/03/2024

experiencias, pretéritas o actuales, y tanto por lo que se refiere a la educación escolar como a la educación social.

Epílogo

Cerramos el artículo con una última reflexión sobre el sentido de enfoques como el ensayado aquí. Nuestro objetivo era hallar posibles coincidencias y complementariedades entre la corriente del Pensamiento Crítico y la Pedagogía Crítica. Para ello partíamos de dos intuiciones. La primera la hemos desarrollado en otro lugar (Trilla, 2007) y, por tanto, ahora no nos extenderemos en ella. Es la intuición de que quizá resultase más productivo, educativamente hablando, esforzarse en buscar aquello que distintas pedagogías comparten que aquello que las separa. Ya sabemos que en la *academia* se propende más a descubrir y subrayar las diferencias que las coincidencias. Y tampoco es que sea desdeñable hacerlo así: en las páginas anteriores, a pesar de que nuestro objetivo no era éste, tampoco hemos escatimado poner en evidencia lo que separa ambas corrientes. Pero a pesar de eso, creemos que debe insistirse en la conveniencia de esforzarse en las coincidencias y complementariedades. Ni que sea porque la educación es siempre una práctica, una acción, y no una omisión -dijera lo que dijera Rousseau-. Y para la acción siempre es mejor sumar que restar: las diferencias restan y las complementariedades suman.

La segunda intuición que ha orientado el artículo es un simple corolario de la primera. Para encontrar coincidencias y complementariedades es menester estar abierto a aquello que, de entrada, nos parece ajeno, distinto, distante o contrario. Encerrarse en el propio paradigma (si es que en pedagogía existen paradigmas de verdad), centrarse en un *centrismo* u otro (sea europeo, americano o de donde sea), declararse ortodoxo de cualquier ortodoxia pedagógica, apuntarse acrítica y sectariamente a cualquier sistema de verdades…; o renunciar de entrada a las aportaciones de la corriente del Pensamiento Crítico porque uno se declara partidario de la Pedagogía Crítica (y viceversa); u oponer a dos gigantes como son Dewey y Freire… quizá no sean las opciones más recomendables para la práctica educativa. Ni tan siquiera, seguramente, para la reflexión pedagógica: ya decía Freire en una de sus frases memorables que "para obtener verdades, es necesario no estar demasiado seguro de ellas" (Freire, P., 1996, p.7).

Recibido: 05/05/2023 Aceptado: 11/03/2024

Referencias bibliográficas

Alumnos de la escuela de Barbiana. (2017). *Carta a una maestra*. PPC Editorial.

Apple, M. W. (1996). *El conocimiento oficial*. Paidós.

Apple, M.W. y Beane, J. A. (Comps.). (2000). *Escuelas democráticas*. Morata.

Ayuste, A. y Trilla. J. (2005). Pedagogías de la modernidad y discursos postmodernos sobre la educación. *Revista de Educación*, *336*, 219-248.

Blair, J. A., (Ed.). (2021). *Studies in Critical Thinking*. Windsor studies in argumentation. https://windsor.scholarsportal.info/omp/index.php/wsia/catalog/book/106

Burbules, N. C. & Berk, N. (1999). Critical Thinking and Critical Pedagogy: Relations, Differences, and Limits. En Popkewitz, T. S. and Fendler, L. (Eds.), *Critical Theories in Education*. (pp. 45-65). Routledge.

Busom, R. (2019). *Kant en Silicon Valley: Pensamiento crítico para emprendedores*. KDP Amazon. Recovered from https://www.amazon.es/Kant-Silicon-Valley-Pensamiento-emprendedores-ebook/dp/B07XWMFB12/ref=asap_bc

Cañas, M., Pinedo, R., y Palacios, A. (2022). Prácticas de aula para promover el pensamiento crítico en el uso de los medios digitales. *Revista de Educación*, *398*, 193-217. https://doi.org/10.4438/1988-592X-RE-2022-398-557

Comas, L.; Valero, A. y Vendrell, M. (2021). Redes sociales. ¿Un asalto al corazón de la democracia? En González, M. R., Igelmo, J. y Jover, G. (Eds.), *Condiciones del pensamiento crítico en el contexto educativo de inicio del siglo XXI* (pp. 93-103). FahrenHouse.

Consejo de la Unión Europea (2018). *Recomendación del consejo de 22 de mayo de 2018 relativa a las competencias clave para el aprendizaje permanente*. Recuperado de https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018H0604(01)&from=SV

Corbett, J. & Guilherme, M. (2021). Critical pedagogy and quality education (UNESCO SDG-4): the legacy of Paulo Freire for language and intercultural communication. *Language and Intercultural Communication*, *21*(4), 447-454. https://doi-org.sire.ub.edu/10.1080/14708477.2021.1962900

Corzo, J. L. y Gesualdi, F. (1992). *Don Milani nella scrittura collettiva*. Edizioni Gruppo Abele.

Recibido: 05/05/2023 Aceptado: 11/03/2024

Corzo Toral, J. L. (2021). La pedagogía de Milani en España e Iberoamérica. *Historia De La Educación*, *39*(1), 337–356. https://doi.org/10.14201/hedu202039337356

Darder, A. (2017). *Freire y Educación*. Morata.

Dewey, J. (1989). *Cómo pensamos. Nueva exposición de la relación entre pensamiento reflexivo y proceso educativo*. Paidós.

Dewey, J. (2004). *Democracia y educación. Una introducción a la filosofía de la educación*. Morata.

Ennis, R. (2005). Pensamiento crítico: un punto de vista racional. *Revista de Psicología y Educación, 1*(1), 47–64.

Ennis, R. (2011). Reflection and perspective. Part I. *Inquiry, 26*(1), 4–18.

Ennis, R. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi, 37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. The California Academia Press. Recovered from https://eric.ed. gov/?id=ED315423

Facione, P. (2020). *Critical Thinking: What It Is and Why It Counts. Measured Reasons LLC*. Recovered from https://www.insightassessment.com/article/critical-thinking-what-it-is-and-why-it-counts

Ferrer y Guardia, F. (1976). *La Escuela Moderna*. Ediciones Jucar.

Fisher, A. (2021). What Critical Thinking Is. In Blair, J. A., (Ed.). *Studies in Critical Thinking* (pp. 7-28). Windsor studies in argumentation. Recovered from https://windsor.scholarsportal.info/omp/index.php/wsia/catalog/ book/106

Freire, P. (1970). *Pedagogía del oprimido*. Tierra Nueva.

Freire, P. (1975). *Las Iglesias, la educación y el proceso de liberaciEón humana en la historia*. La Aurora.

Freire, P. (1984). *La importancia de leer y el proceso de liberación.* Siglo XXI.

Freire, P. (1985). *¿Extensión o comunicación? La concientización en el medio rural*. Siglo XXI.

Freire, P. (1996). Prólogo. En Gadotti, M. *Pedagogía de la praxis* (pp. 6-8). Miño y Dávila Editores.

Gadotti, M. (1996). *Pedagogía de la praxis*. Miño y Dávila Editores.

Giroux, H. A. (1990). *Los professores como intelectuales. Hacia una pedagogia crítica del aprendizaje*. Paidós.

Giroux, H. A. (2020). *On Critical Pedagogy*. Bloomsbury Academic.

Giroux, H. A. (2022). Terrorismo pedagógico y esperanza en la era de la política fascista. En Rodríguez Martínez, C. y Imbernón Muñoz, F.

(Eds.), *De las políticas educativas a las prácticas escolares* (pp. 199-212). Morata.

Ichikawa, H. (2020). A theory of hope in critical pedagogy: An interpretation of Henry Giroux, *Educational Philosophy and Theory, 54*(4)*, 384-394*. https://doi.org/10.1080/00131857. 2020.1840973

Innerarity, D. (2022). *La sociedad del desconocimiento*. Galaxia Gutenberb.

Kincheloe, J. L. (2004). La Pedagogía del oprimido y el papel central desempeñado por Paulo Freire en mi vida. En Araujo Freire, A. M. (Coord.), *La pedagogía de la liberación en Paulo Freire* (pp. 151-166). Graó.

Kincheloe, J. L. (2008). La Pedagogía Crítica en el siglo XXI: Evolucionar para sobrevivir. En MCLaren, P. y Kincheloe, J.L. (Eds.), *Pedagogía Crítica. De qué hablamos, dónde estamos* (pp. 25-69). Graó.

Kincheloe, J. L., MacLaren, P. y Steinberg, S. R. (2003). Palabras preliminares. En Giroux, H. A. *Pedagogía y política de la esperanza* (pp. 11-18). Amorrortu.

Lipman, M. (1998). *Pensamiento complejo y educación*. Ediciones de la Torre.

López, G. (2012). Pensamiento crítico en el aula. *Docencia e investigación*, Año 37, 22, 41-60.

Macedo, D. (2000). Una pedagogía antimétodo. Una perspectiva freiriana. *Educación XX1*, (3), 53-61.

McLaren, P. (2022). Pedagogía crítica y lucha de clases en la era del terror neoliberal. En Rodríguez Martínez, C., y Imbernón Muñoz, F. (Eds.), *De las políticas educativas a las prácticas escolares* (pp. 231-282). Morata.

McLaren, P. y Kincheloe, J. L. (Eds.). (2008). *Pedagogía Crítica. De qué hablamos, dónde estamos*. Graó.

Paul, R. (1981). Teaching critical Thinking in the "Strong" Sense: A Focus On Self-Deception, World Views, and a Dialectical Mode of Analysis. *Informal Logic*, *4*(2)*,* 2-7. https://doi.org/10.22329/il.v4i2.2766

Paul, R., & Elder, L. (2003). *La mini-guía para el Pensamiento Crítico. Conceptos y Herramientas*. Fundación para el Pensamiento Critico.

Paul, R. & Elder, L. (2007). *A guide for educators to critical thinking competency standards*. Dillon Beach, CA; Foundation for Critical Thinking.

Saiz, C. & Rivas, S. F. (2008). Evaluación del pensamiento crítico: una propuesta para diferenciar formas de pensar. *Ergo, Nueva Época*, *22-23*, 25-66.

Recibido: 05/05/2023 Aceptado: 11/03/2024

Trilla, J. (1989). La relación teoría-práctica en la pedagogía de Paulo Freire. *Temps d'Educació, 1,* 243-254.

Trilla, J. (1993). *Otras educaciones. Animación sociocultural, formación de adultos y ciudad educadora*. Anthropos.

Trilla J. (2007). Educational Discourse and Educational Practice. *Encounters on Education*, *8*, 127–142.

Wheeler-Bell, Q. (2019). An immanent critique of critical pedagogy. *Educational Theory, 69*(3), 265–281. https://doi.org/10.1111/edth.12368

Información de contacto: Ana Ayuste González. Institut de Recerca en Educació. Universitat de Barcelona. Passeig Vall d'Hebron, 171, 08035 Barcelona. E-mail: anaayuste@ub.edu

La experiencia académica del alumnado gitano en la universidad

The academic experience of Roma students at university

https://doi.org/10.4438/1988-592X-RE-2024-406-648

Silvia Abad Merino
https://orcid.org/0000-0003-1132-3051
Universidad de Córdoba

Fernando Macías-Aranda
https://orcid.org/0000-0002-1569-6659
Universidad de Barcelona

Olga Serradell Pumareda
https://orcid.org/0000-0003-4077-1400
Universidad Autónoma de Barcelona

Ainhoa Flecha
https://orcid.org/0000-0001-6910-9715
Universidad Autónoma de Barcelona

Resumen

El Pueblo Gitano, a pesar de ser la principal minoría étnica europea con más de 600 años en el continente, sigue enfrentando discriminación y exclusión, impactando particularmente en sus logros educativos. Solo el 1-4% de los gitanos tiene educación universitaria, contrastando con el 44% de otros grupos. Asimismo, los gitanos que alcanzan el sistema universitario tienen más probabilidades de abandonar sus estudios, debido a que suelen ser estudiantes no tradicionales y/o de primera generación, y a múltiples formas de discriminación, como sus ingresos,

Recibido: 01/09/2023 Aceptado: 06/06/2024

y con un especial riesgo para las mujeres gitanas que llegan a la universidad. Este artículo se basa en los resultados del proyecto I+D "UNIROMA" (2019-2022), centrado en la población gitana de las universidades españolas. Utilizando la Metodología Comunicativa, la investigación adoptó un diseño secuencial mixto con dos fases. En la primera fase cuantitativa obtuvo 98 respuestas de estudiantes gitanos universitarios. La segunda fase se centró en datos cualitativos, con 20 historias de vida a estudiantes gitanos universitarios y 15 entrevistas a expertos nacionales e internacionales. Este artículo recoge resultados sobre la información sociodemográfica y la experiencia universitaria de estos estudiantes. Los resultados sugieren que el alumnado gitano universitario enfrenta dificultades relacionadas con la configuración de su hogar, su situación laboral o sus recursos económicos, afectando su experiencia universitaria. Equilibrar estudios, trabajo y familia influye en su dedicación académica, al igual que en estudiantes de minorías étnicas y primera generación. Los exámenes generan estrés y dominar lenguas extranjeras es un desafío común. La discusión destaca la necesidad de seguir desarrollando medidas, especialmente desde la universidad, para asegurar la permanencia de los estudiantes gitanos, en su mayoría de primera generación, para que completen con éxito la educación sus estudios universitarios.

Palabras clave: universidad, Pueblo Gitano, experiencia académica, estudiantes no tradicionales, estudiantes de primera generación.

Abstract

Roma People, despite being the main European ethnic minority with over 600 years on the continent, continue to face discrimination and exclusion, particularly impacting their educational achievements. Only 1-4% of Roma have university education, contrasting with 44% of other groups. Likewise, Roma individuals who enter the university system are more likely to drop out due to being non-traditional and/or first-generation students, and facing multiple forms of discrimination such as income, with a particular risk for Roma women attending university. This article is based on the results of the R&D project "UNIROMA" (2019-2022), focused on the Roma population in Spanish universities. Using Communicative Methodology, this research adopted a mixed sequential design with two phases. In the first quantitative phase, 98 responses were obtained from Roma university students. The second phase focused on qualitative data, including 20 life stories of Roma university students and 15 interviews with national and international experts. This article presents findings on sociodemographic information and the university experience of these students. Results suggest that Roma university students face difficulties related to their home environment, employment situation, and economic resources, impacting their university experience. Balancing studies, work, and family influences their academic dedication, as observed in ethnic minority and first-generation students. Exams create stress

Recibido: 01/09/2023 Aceptado: 06/06/2024

and mastering foreign languages is a common challenge. Discussion highlights the need to continue developing measures, especially within the university, to ensure the retention of Roma students, mostly first-generation, to successfully complete their university education.

Keywords: university, Roma People, academic experience, non-traditional students, first-generation students.

Introducción

La historia del pueblo gitano ha estado tradicionalmente asociada a la marginalización, la pobreza y la exclusión socioeducativa. Aunque las recientes iniciativas internacionales han favorecido importantes logros en relación con la alfabetización y la escolarización obligatoria de las comunidades romaníes (Comisión Europea, 2020), los estudiantes gitanos continúan liderando las tasas más bajas de éxito académico en el sistema educativo (Agencia de los Derechos Fundamentales de la Unión Europea EU-FRA, 2022; De la Rica et al., 2019; Rutigliano, 2020).

En el contexto universitario, los datos disponibles hasta 2018 indican que en los países de la OCDE (2020), el 44.4% de los adultos hasta los 34 años alcanzaron la educación superior. Sin embargo, se estima que la población romaní con estudios universitarios se encuentra entre el 1% y el 4% (The Velux Foundations, 2019). Concretamente, las bajas tasas de graduación de la población gitana en las universidades españolas reflejan la necesidad urgente de estudiar las dificultades que afrontan los estudiantes gitanos, así como las acciones basadas en la evidencia que contribuyen a superarlas.

En los últimos años, se han efectuado importantes esfuerzos en educación primaria y secundaria para aumentar la población gitana que accede a la educación superior (Brüggemann, 2014; Flecha, 2018). No obstante, estas medidas necesitan incluir aquellos mecanismos que mitiguen las dificultades que enfrentan los estudiantes gitanos durante su experiencia académica universitaria para aumentar efectivamente las tasas de graduación terciaria. Recientemente, se han comenzado a implementar actuaciones orientadas a garantizar la equidad y la inclusión de los estudiantes gitanos en la educación superior. Sin embargo, el estudio longitudinal de la efectividad de estas iniciativas y de las barreras

estructurales que enfrenta la población romaní en la universidad aún está insuficientemente desarrollado (Garaz y Torotcoi, 2017; Rutigliano, 2020).

El presente estudio forma parte del proyecto I+D UNIROMA (2019-2022), financiado por el Ministerio de Ciencia, Innovación y Universidades, el cual tuvo como objetivo aportar evidencias científicas que contribuyan al éxito del alumnado gitano en la universidad, promoviendo así la inclusión social del grupo étnico más representativo de España y Europa. Por consiguiente, el proyecto contribuyó especialmente al Objetivo 4 de Desarrollo Sostenible, que aboga por una educación inclusiva, equitativa y de calidad para todos.

En el ámbito de la Educación Superior, las diferencias en el éxito académico de los estudiantes de grupos étnicos minoritarios respecto a sus iguales del grupo social dominante empiezan a definirse desde los condicionantes de las etapas educativas anteriores (Lens y Levrau, 2020). Concretamente, se documentan la discriminación experimentada por el alumnado romaní en el sistema educativo y las dificultades que este enfrenta desde los niveles escolares más básicos (Lambrev, 2020).

Desde la primera infancia, los niños romaníes enfrentan las limitaciones de un sistema educativo estandarizado y monocultural. La falta de maestros familiarizados con su cultura dificulta la comunicación efectiva con ellos y sus familias (Bîrle et al., 2014). Estos desafíos interactúan con discriminación (Eurofound, 2016; Kende et al. 2020), pobreza (Kamberi et al., 2015) y factores socioculturales que afectan su desenvolvimiento en el sistema educativo (Mulcahy et al. 2017), resultando en bajo rendimiento, absentismo y alto abandono escolar entre los romaníes.

En la Unión Europea, las políticas sociales y educativas muestran barreras de acceso y participación que generan segregación y baja calidad educativa para los romaníes (Alexiadou, 2019). La discriminación institucionalizada perjudica a los estudiantes minoritarios (Brüggemann y D'Arcy, 2017). Las familias desconectadas suelen mantener a sus hijos en escuelas gueto y estigmatizadas (Neumann, 2016). Los diagnósticos erróneos debido a pruebas estandarizadas y la vulnerabilidad socioeconómica los colocan en aulas especiales de baja calidad (Amnesty International y European Roma Rights Center, 2017; Cashman, 2017).

Organismos oficiales confirman que los niños romaníes a menudo están en escuelas de baja calidad, en niveles educativos inferiores y aulas segregadas (Agencia de los Derechos Fundamentales de la Unión Europea EU-FRA, 2016). La segregación y la asignación a grupos de compensación

Recibido: 01/09/2023 Aceptado: 06/06/2024

educativa sin garantías agravan la brecha educativa (CREA, 2010; Parra Toro et al., 2017). Estudiar en aulas segregadas afecta la identidad y autoestima, y limita las oportunidades de permanencia escolar (Messing, 2017).

Los estudiantes de minorías étnicas afrontan obstáculos en el sistema educativo que los colocan en desventaja en relación con el grupo dominante en términos de acceso a la educación superior, debido a disparidades en las calificaciones de ingreso universitario. Por ejemplo, en el Reino Unido, las posibilidades de que estudiantes étnicos minoritarios obtengan un título de calidad son la mitad en comparación con estudiantes blancos (Richardson, 2018). Similarmente, en España los estudiantes romaníes están significativamente por debajo de la representación promedio en la educación superior en comparación con el grupo mayoritario (Ministerio de Derechos Sociales y Agenda 2030, 2021). Este dato afecta especialmente a campos como ciencias y tecnología, como su análisis también indica en el conjunto de Europa del Este, lo que limita sus perspectivas laborales (Garaz y Torotcoi, 2017). Esto destaca la necesidad de examinar cómo las universidades contribuyen a la injusticia social y de implementar acciones para apoyar a los estudiantes gitanos a través de compromiso académico y político (Arday y Mirza, 2018).

Investigaciones recientes enfatizan que las medidas efectivas parten del conocimiento de los estudiantes y la comprensión de las barreras interseccionales que enfrentan como miembros de minorías étnicas, estudiantes no tradicionales y de primera generación, además de tener situaciones socioeconómicas desfavorables (Flecha et al., 2022). La discriminación y la exclusión que experimentan los estudiantes gitanos durante sus estudios también interactúan con su aprendizaje sobre la dinámica universitaria y su identidad personal (McDonagh y Fonseca, 2022). Algunos ocultan su identidad por miedo a ser excluidos, mientras que otros la abrazan para normalizar el éxito educativo gitano y contribuir a su comunidad (Hinton-Smith y Padilla-Carmona, 2021; Hinton-Smith et al., 2017).

Investigaciones en las últimas dos décadas muestran el deterioro de la salud mental y el estrés en estudiantes de minorías étnicas y primera generación (Ogunyemi et al., 2019), impactando negativamente su experiencia de aprendizaje. Esto está relacionado con las barreras de acceso, el agotamiento por enfrentar discriminación, demostrar valía y cumplir expectativas (Kundu, 2019). La presión estereotipada también afecta a los gitanos académicamente, ya que resaltar su identidad durante

tareas académicas disminuye su rendimiento (Slobodnikova y Randolph-Seng, 2021; Želinský, 2022).

A pesar de los programas de preparación en universidades, la discriminación institucional persiste para estudiantes de minorías étnicas. Muchos sienten la necesidad de educar a otros sobre raza y justicia social, enfrentan prejuicios raciales y se sienten aislados en la universidad (Harris y Linder, 2018). La percepción de que sus contribuciones son ignoradas desalienta a los estudiantes subrepresentados y genera desconfianza (Allaire, 2019).

La intersección de origen étnico y clase social incrementa las dificultades para manejar discriminación y marginación en la universidad (Silver, 2019). Esto es más pronunciado entre estudiantes gitanos, y su interacción con la etnia predice actitudes de prejuicio (Urbiola et al., 2022). Sin embargo, la movilidad internacional puede ser transformadora para estudiantes gitanos, permitiéndoles enfrentar la discriminación y avanzar profesionalmente (Marcu, 2020; Roberts, 2021).

A pesar de las dificultades, la investigación en España sobre estudiantes gitanos en la universidad es limitada. Este estudio analiza perfiles de estudiantes gitanos y sus desafíos académicos, debido a la rigidez universitaria hacia estudiantes no tradicionales y de primera generación. Es decir, hacia estudiantes que ingresan en la universidad a una edad posterior, tienen cargas socioeconómicas o son los primeros en sus familias que estudian en la universidad. Concretamente, el Centro Nacional de Estadísticas de Educación (NCES) de los Estados Unidos (2024) concreta que un estudiante no tradicional presenta alguna o varias de las siguientes características: ingreso en la universidad posterior a la norma; matrícula a tiempo parcial; independencia financiera; trabajo a tiempo completo; con personas dependientes distintas de su cónyuge; modelo unifamiliar; o acceso sin haber obtenido el título de la escuela secundaria estándar.

Diseño metodológico

El proyecto UNIROMA se basó en la Metodología Comunicativa que considera a las personas como agentes sociales con lenguaje y capacidad de acción, promoviendo el diálogo entre investigadores, participantes y otros actores involucrados en las realidades estudiadas. Este enfoque metodológico permite describir la realidad y contribuir a su

Recibido: 01/09/2023 Aceptado: 06/06/2024

transformación a través del impacto social, político y científico generado por la investigación (Gómez et al., 2010).

En esta investigación, se otorgó especial relevancia a la participación activa de los agentes sociales en todas las etapas del proceso, desde el diseño inicial hasta la formulación de conclusiones. Para ello, se constituyó un Consejo Asesor integrado por personas gitanas que cursaban estudios universitarios, representantes de asociaciones gitanas y personas con experiencia en gestión universitaria. Su valiosa contribución fue crucial para el diseño del trabajo de campo, la recopilación de datos y la identificación de perfiles de estudiantes universitarios.

La investigación se desarrolló mediante un diseño de método mixto secuencial que incluyó dos fases: una fase inicial de recopilación de datos cuantitativos, en la que participaron estudiantes gitanos universitarios, y una segunda fase de recopilación de datos cualitativos, donde participaron tanto estudiantes gitanos universitarios como personas expertas en el ámbito del estudio.

Para abordar la falta de investigaciones sobre personas gitanas cursando estudios universitarios en el país, se llevó a cabo una fase inicial exploratoria. Esta incluyó una revisión exhaustiva de la literatura y la administración de un cuestionario en línea a personas gitanas matriculadas en estudios universitarios en ese momento. Los datos obtenidos permitieron identificar diferentes perfiles de estudiantes universitarios, que fueron tenidos en cuenta para seleccionar la muestra del estudio cualitativo.

El trabajo de campo cuantitativo desarrollado durante la primera fase de la investigación, se realizó durante un estudio longitudinal que incluyó dos etapas de recogida de datos. La primera etapa se realizó entre mayo y julio de 2020 y la segunda tuvo lugar entre abril y mayo de 2022, obteniendo 98 y 51 respuestas válidas, respectivamente. Los perfiles de los estudiantes se seleccionaron a partir de las respuestas proporcionadas en el cuestionario de la primera fase del proyecto.

En la segunda fase de la investigación se desarrolló el trabajo de campo cualitativo, donde se recogieron 20 historias de vida elaboradas por estudiantes gitanos de diversas universidades y ámbitos académicos en España. Además, en esta fase se realizaron entrevistas en profundidad a 16 expertos nacionales e internacionales con experiencia en gestión universitaria y programas de apoyo a estudiantes universitarios de minorías étnicas y primera generación. Los resultados permitieron

elaborar recomendaciones para el apoyo a estudiantes romanís en la universidad (Serradell et al., 2022).

Muestra

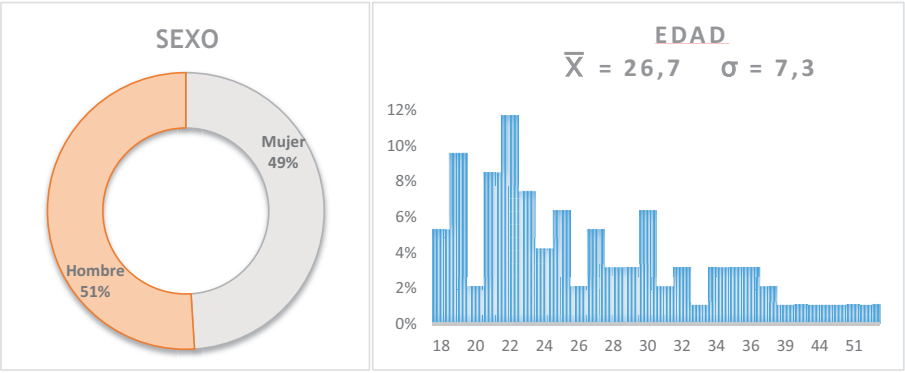
En la fase cuantitativa del estudio, la falta de datos étnicos en España impidió realizar un muestreo probabilístico en esta área. En su lugar, se optó por utilizar un método de muestreo de bola de nieve para identificar los perfiles de estudiantes gitanos en las universidades españolas. Inicialmente, el reclutamiento de participantes se realizó a través de redes establecidas con asociaciones gitanas y se completó a través de los contactos que estas proporcionaron, con el objetivo de llegar a más individuos.

Debido a que la población gitana en España es relativamente pequeña, con aproximadamente 1.000.000 personas gitanas en el país y solo alrededor del 1-2% de ellos cursando estudios universitarios, obtener respuestas de aquellos que cumplían con el perfil analizado en la investigación (personas gitanas matriculadas en estudios universitarios) presentó dificultades notables, lo que constituye una de las principales limitaciones del estudio (Schanze y Zins, 2019; Salamonska y Czeranowska, 2018). Algunas posibilidades para aumentar el acceso a la población gitana universitaria pueden ser la inclusión de más personas gitanas en el equipo investigador y/o de trabajo, la realización de actos y reuniones en la universidad dirigidos a estudiantes que pertenezcan a minorías culturales y/o religiosas o la participación de más organizaciones gitanas de la sociedad civil, especialmente las que trabajan con proyectos educativos.

La segunda dificultad, como es común en estudios longitudinales, fue asegurarse que las personas que completaron el cuestionario inicial también completaran el cuestionario de seguimiento, es decir, el que se debía realizar durante la segunda vuelta. La muestra participante del cuestionario inicial incluyó a 98 personas gitanas que, en el momento de la encuesta, estaban matriculadas en programas de estudios regulares en diversas universidades españolas. Se obtuvieron respuestas de 33 universidades y 12 comunidades autónomas del país, destacando Cataluña (42%), Andalucía (22%) y la Comunidad Valenciana (10%) entre ellas.

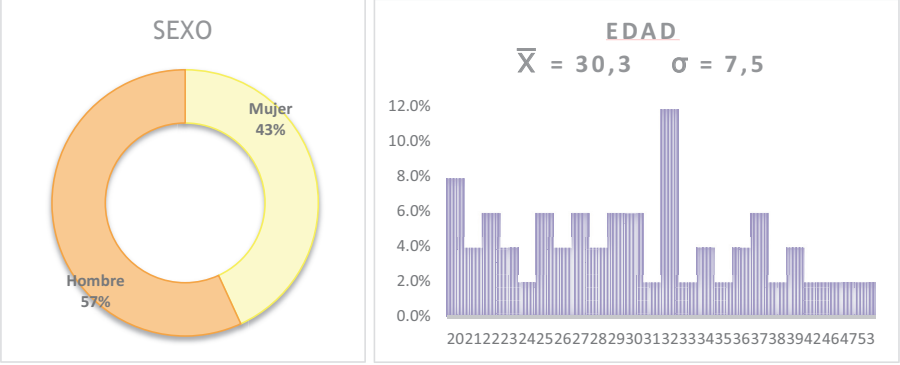
Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO I. Distribución de la muestra por sexo y edad (primera vuelta, fase cuantitativa)



Fuente: elaboración propia.

GRÁFICO II. Distribución de la muestra por sexo y edad (segunda vuelta, fase cuantitativa)



Fuente: elaboración propia.

En cuanto a la distribución por edad y sexo (ver Gráfico I), la muestra está formada por un 51% de hombres y un 49% de mujeres, con una media de edad de 26,7 años (desviación típica, 7,3 años) sin diferencias significativas por sexo (edad media entre las mujeres, 26,28; y edad media entre los hombres, 27,06).

En cuanto a la segunda vuelta (ver Gráfico II), se obtuvieron 51 respuestas de 9 Comunidades Autónomas, predominando las provenientes de Cataluña (52,9%) y Comunidad Valenciana (23,5%). En esta ocasión, la distribución por sexo reflejó una mayor representación de hombres, con un 57% frente al 43% de mujeres, y una media de 30,3 años con una desviación típica de 7,5 años.

En relación con la fase cualitativa del trabajo de campo, se recogieron 20 historias de vida elaboradas por estudiantes gitanos que se encontraban estudiando distintas titulaciones en diferentes universidades españolas (ver Tabla I). La muestra estuvo formada por 10 mujeres y 10 hombres, de los cuales 11 residían en Cataluña y 9 en Andalucía. Los perfiles de los estudiantes fueron seleccionados utilizando las respuestas proporcionadas por 98 estudiantes universitarios gitanos en el cuestionario de la primera fase del proyecto. De esta manera, las 20 historias de vida reflejan la diversidad en términos de género, perfiles académicos y áreas de conocimiento, incluyendo ciencias sociales y humanidades, ciencias de la salud y ciencias naturales. La mayoría de los participantes entrevistados (9 mujeres y 8 hombres) estaban matriculados en programas de grado, mientras que 2 hombres y 1 mujer estaban estudiando un máster. Además, 11 participantes presentaban un perfil de estudiante tradicional, mientras que los 9 restantes tenían un perfil no tradicional.

Las vías mayoritarias de acceso a la universidad son las más frecuentes, como el bachillerato y la realización de las Pruebas de Acceso a la Universidad (PAU) o la Formación Profesional Superior con o sin las PAU. Sin embargo, las minoritarias hacen referencia a las menos comunes,

TABLA I. Perfiles de la muestra (fase cualitativa)

	Cataluña		**Andalucía**	
	Mujeres	Hombres	Mujeres	Hombres
Estudiante tradicional 17-24 años, acceso por vías mayoritarias sin responsabilidades laborales o familiares.	**3**	**2**	**3**	**3**
Estudiante no tradicional	**2**	**4**	**2**	**1**
17-24 años, acceso por vías mayoritarias con responsabilidades laborales o familiares.	1	1	1	0
>25 años, acceso por vías minoritarias.	1	3	1	1

Fuente: elaboración propia.

Recibido: 01/09/2023 Aceptado: 06/06/2024

como la prueba de acceso a la universidad para mayores de 25 y 45 años. La selección de participantes consideró criterios que permitieran incluir los diversos perfiles identificados en el cuestionario. Es decir, asegurar la variedad de áreas de estudio, equilibrar la representación geográfica con mayor población gitana, incorporar estudiantes de postgrado sin afectar la realidad mayoritaria de estudiantes de grado, garantizar la participación de estudiantes con becas y aquellos que no estuvieran asociados con el movimiento gitano. De esta manera, logramos reflejar la diversidad de trayectorias y obtener información relevante y significativa.

Una de las principales dificultades que enfrentamos fue identificar estudiantes en disciplinas STEM, especialmente en el ámbito de la ingeniería. Para establecer contacto con los estudiantes, contamos con la colaboración de profesores universitarios y asociaciones gitanas.

Instrumentos

Para llevar a cabo el estudio longitudinal de la fase cuantitativa, se han desarrollado dos instrumentos en formato online: un cuestionario inicial y un cuestionario de seguimiento. En el diseño del cuestionario inicial, el equipo de investigación colaboró con el Consejo Asesor y, especialmente, con CampusRom - Red Gitana Universitaria, una asociación que agrupa a estudiantes romaníes. Se utilizaron como base cuestionarios previos sobre trayectorias universitarias en general y estudiantes de minorías étnicas en particular, como el Student Experience Survey (Times Higher Education, 2018), el cuestionario de caracterización psicopedagógica del estudiante universitario (Márquez y Ordaz, 2014), el cuestionario CEVEAPEU (Gargallo et al., 2009), el NSSE (Center for Postsecondary Research, 2019) y el First Generation College Survey (Students for Education Reform, 2017).

El cuestionario inicial se validó a través de un pilotaje para garantizar su viabilidad, confiabilidad y validez, así como para ajustar el tiempo necesario para completarlo. Después de realizar las modificaciones pertinentes, se generó el cuestionario final, que consta de 86 ítems y los siguientes bloques temáticos:

- Trayectoria académica previa y vías de acceso;
- Compaginación con otras actividades;

- Estudios actuales: ámbito académico, interacción social y reconocimiento cultural;
- Dificultades;
- Medidas;
- Preguntas sociodemográficas.

El cuestionario de seguimiento se centró en conocer el progreso de los estudiantes entrevistados inicialmente en relación con sus estudios universitarios, incluyendo si aún estaban estudiando, si se habían graduado o si habían abandonado, y cuáles fueron las principales dificultades que enfrentaron.

Para llevar a cabo la fase cualitativa se utilizaron historias de vida realizadas por estudiantes universitarios gitanos y se realizaron entrevistas en profundidad a personas expertas.

Las historias de vida abordaron los siguientes temas:

- Trayectoria educativa previa a la universidad;
- Vía de acceso a la universidad;
- Dificultades burocráticas en la universidad;
- Dificultades académicas en la universidad;
- Las relaciones sociales en la universidad.

En las entrevistas con las personas expertas se abordaron aspectos relacionados con 1) Programas de apoyo, acompañamiento e información; 2) Ayuda financiera; y 3) Participación social. Todas estas medidas se encontraron vinculadas a estudiantes de minorías étnicas y/o no tradicionales (segunda generación, mayores de 25 años, responsabilidades laborales y/o familiares).

El presente artículo muestra los resultados de la fase cuantitativa y cualitativa del estudio, vinculados a la información sociodemográfica y académica de los estudiantes. Concretamente, el análisis incluye:

- Información sociodemográfica
  - Forma de convivencia y datos del hogar
  - Situación laboral y económica
  - Recursos para el estudio
  - Capital cultural familiar
- Experiencia universitaria
  - Dedicación académica
  - Autosuficiencia

Recibido: 01/09/2023 Aceptado: 06/06/2024

Resultados

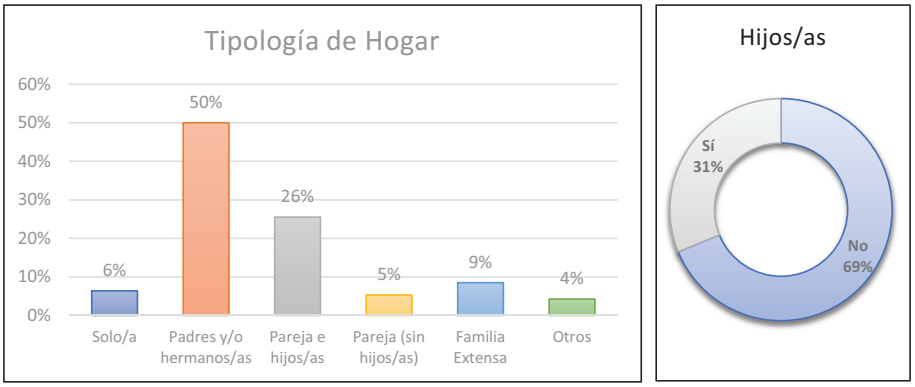
Los resultados obtenidos muestran información sociodemográfica y académica relacionada con las barreras que experimentan los estudiantes gitanos universitarios en su trayectoria.

Información sociodemográfica

Forma de convivencia y datos del hogar

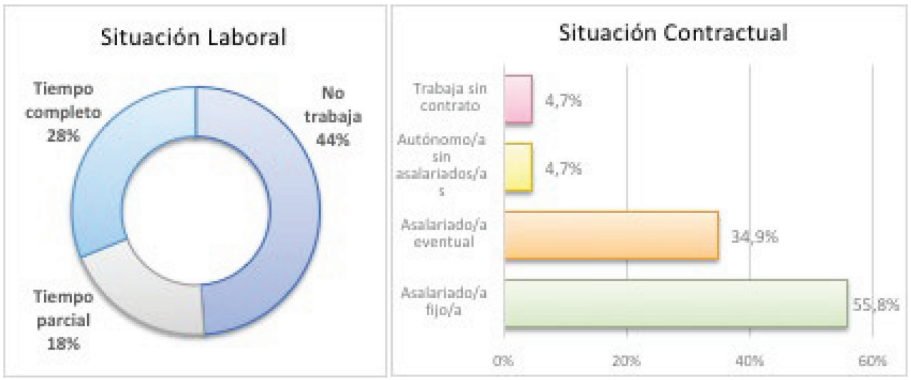
La mayoría de los estudiantes de la muestra viven con sus padres y/o hermanos (50%). No obstante, la proporción de estudiantes emancipados alcanza casi la mitad de la muestra, considerando que el 31% vive en pareja (26% con hijos y 5% sólo con la pareja) y un 6% vive solo. Teniendo en cuenta que el 26% de la muestra tiene hijos, es relevante constatar que las responsabilidades familiares constituyen una dificultad para el seguimiento de estudios tanto por la dedicación horaria que suponen como porque aumenta la necesidad de generar recursos económicos propios. Además, los participantes viven en unidades familiares grandes, con un promedio de 4,26 miembros y una desviación estándar de 1,73 (ver Gráfico III).

GRÁFICO III. Tipología de hogar e hijos/as (%)



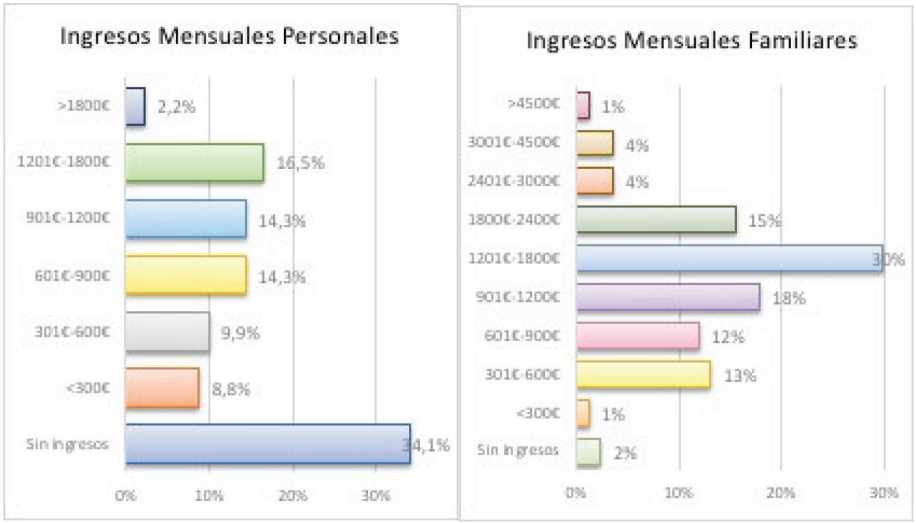
Fuente: elaboración propia.

GRÁFICO IV. Situación laboral y contractual (%)



Fuente: elaboración propia.

GRÁFICO V. Ingresos mensuales personales y familiares (%)



Fuente: elaboración propia.

Recibido: 01/09/2023 Aceptado: 06/06/2024

Situación laboral y económica

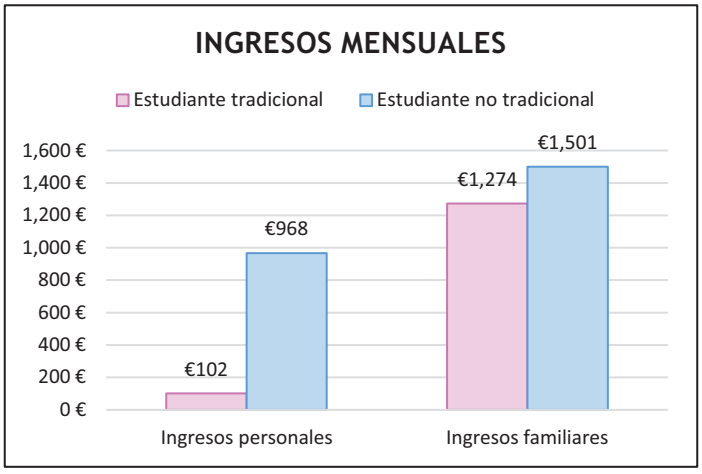
El 46% de las personas participantes trabaja, bien sea a tiempo parcial (18%) o a tiempo completo (28%). Respecto a la situación contractual, el 56% tienen contrato indefinido mientras que un 35% tiene contrato temporal. A estos, habría que añadir casi un 5% de autónomos y el mismo porcentaje de personas que trabajan sin contrato (ver Gráfico IV).

El análisis de los datos relativos a los ingresos arroja una situación de precariedad económica, con un promedio de ingresos familiares de 1411€ (con una desviación de 924€), que cabe poner en relación con el tamaño promedio del hogar que frecuentemente supera los 4 miembros (ver Gráfico V).

Los estudiantes entrevistados destacan la precariedad económica y el componente de estrés asociado a la falta de seguridad de unos ingresos y una vivienda estable.

> De diez asignaturas me han quedado tres y me están costando, porque tenemos un estilo de vida que es más humilde. Siempre tenemos para tirar, con el Primo o con quien sea. Hemos tenido dos o tres desahucios

GRÁFICO VI. Ingresos mensuales según perfil de estudiante (€)



Fuente: elaboración propia.

Recibido: 01/09/2023 Aceptado: 06/06/2024

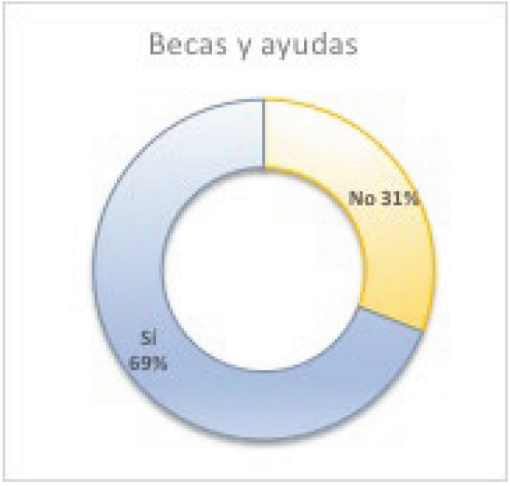
> porque no tenemos para pagar el alquiler y todo esto repercute, es más complicado que otra persona que tiene su casa y sus ingresos… (José, estudiante no tradicional, ciencias sociales y humanidades, Andalucía).

El análisis según el perfil de estudiante, los estudiantes no tradicionales, al ser mayores y trabajar más horas, disponen de ingresos significativamente más altos que los estudiantes tradicionales (ver Gráfico VI).

En consonancia con esta situación económica, los datos muestran que el 69% recibe algún tipo de beca o apoyo al estudio (ver Gráfico VII). Estos datos son coherentes con la posición socioeconómica de la población gitana en España, lo que representa una barrera para el seguimiento de estudios postobligatorios. Además, la percepción de la beca general del Ministerio suele suponer para los estudiantes universitarios una presión añadida. Generalmente, los estrictos criterios académicos que se imponen para mantenerla pueden ser de difícil cumplimiento para aquellos estudiantes que tienen que compaginar los estudios con otras obligaciones laborales y/o familiares.

Concretamente, la percepción de los estudiantes en relación con los rígidos criterios para la obtención de la beca general muestra las

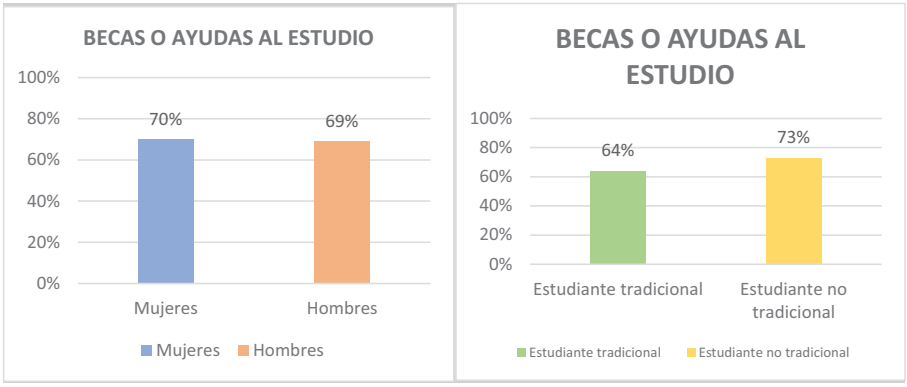
GRÁFICO VII. Beneficiarios/as de algún tipo de beca y/o ayuda (%)



Fuente: elaboración propia.

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO VIII. Becas/Ayudas por sexo y perfil de estudiante (%)



Fuente: elaboración propia.

dificultades que implica para aquellas personas con circunstancias vitales más desfavorables.

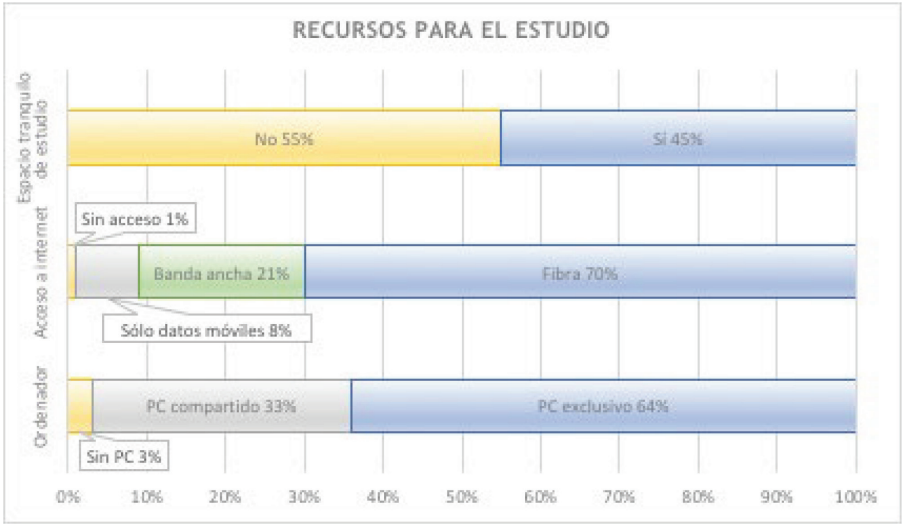
> He contado con becas por suerte durante muchas veces y mis padres me intentan ayudar en todo lo que pueden, trabajar y sacar buenas notas es muy difícil. No me da la vida. (Rocío, estudiante no tradicional, ciencias de la salud, Catalunya).

El porcentaje de estudiantes que recibe algún tipo de beca o ayuda al estudio es más alto entre los estudiantes tradicionales (73%) que entre los no tradicionales (64%). Sin embargo, no existen diferencias al respecto entre hombres y mujeres (ver Gráfico VIII).

Recursos para el estudio

El análisis de los recursos para el estudio en los hogares de los participantes arroja también resultados interesantes, más aún teniendo en cuenta que la encuesta inicial fue realizada entre mayo y junio de 2020, es decir, en el periodo en que las universidades desplazaron toda su actividad académica a formato *online* como consecuencia de

GRÁFICO IX. Recursos para el estudio (%)



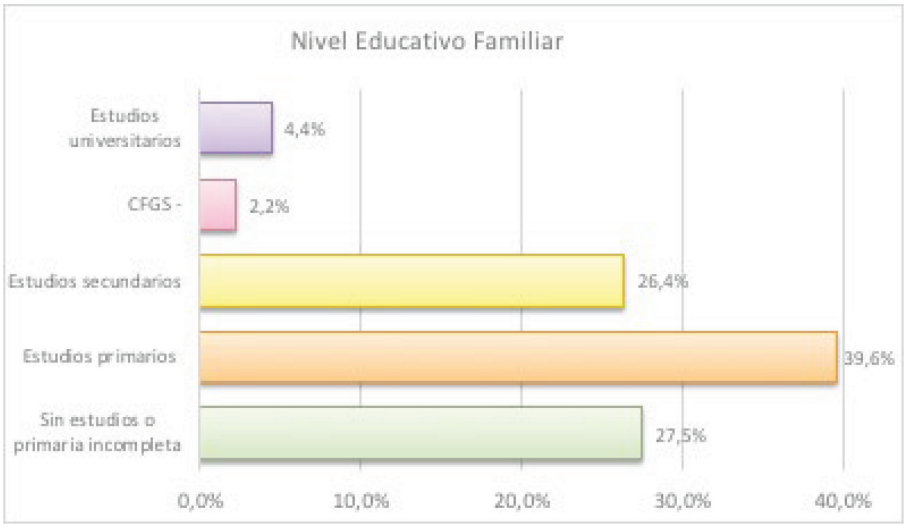
Fuente: elaboración propia.

las medidas adoptadas para hacer frente a la pandemia del COVID19. Específicamente, más de la mitad de los estudiantes (55%) no disponía de un espacio de estudio tranquilo, un 33% no tenía ordenador de uso individual y un 3% no tenía ordenador en casa. En relación con la conexión a internet, un 8% sólo disponía de datos móviles y un 1% no tenía acceso a internet (ver Gráfico IX).

La creciente importancia de la actividad académica virtual que se experimentó durante los últimos años y que se manifestó muy especialmente como consecuencia de las medidas relacionadas con la pandemia, desplazando buena parte de la formación universitaria a formatos virtuales o híbridos, ha supuesto una mayor importancia de los recursos tecnológicos y el espacio para el seguimiento de estudios. Los efectos de esta situación se han visto especialmente reflejados en las poblaciones más vulnerables y el aumento de la brecha educativa respecto a la población mayoritaria.

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO X. Nivel educativo familiar[1] (%)



Fuente: elaboración propia.

Capital cultural familiar

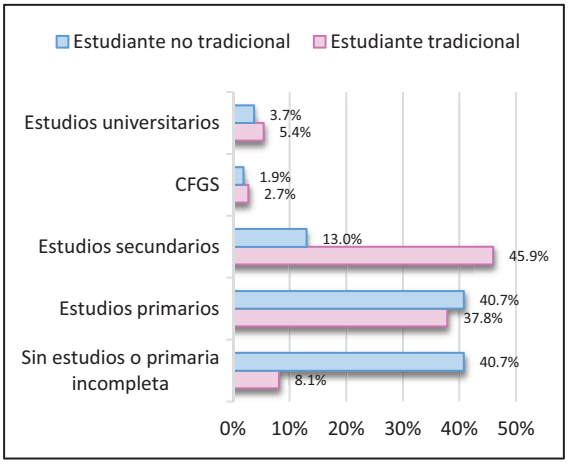
El estudio de la presencia de minorías étnicas en la universidad se aborda frecuentemente ligado a los estudiantes de primera generación, es decir, aquellos que son los primeros de sus familias en alcanzar la educación superior. En grupos culturales con poca tradición académica es frecuente que la mayoría de estudiantes sean de primera generación, lo cual constituye una doble fuente de discriminación y de dificultad en la trayectoria académica.

Este es el caso de la población gitana universitaria, tal y como queda reflejado en el nivel educativo familiar de las personas que han participado en el estudio. En concreto, el trabajo de campo cuantitativo refleja que el 67% de las familias tiene como máximo, estudios primarios, el 26% secundarios, el 2% estudios de Grado Superior, y tan sólo en un 4% de los casos, alguno de los progenitores tiene estudios universitarios (ver Gráfico X).

[1] Para determinar el nivel educativo familiar se ha cogido, en cada caso, el del progenitor con un nivel educativo más alto.

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XI. Nivel educativo familiar según perfil de estudiante (%)



Fuente: elaboración propia.

> Mis padres tienen, digamos, la enseñanza básica, no tienen el título de secundaria y demás, y no llegaron a los estudios superiores y yo soy la primera de mis hermanos. Así, pues la primera que accedo a la universidad… (Rosa, estudiante tradicional, ciencias sociales y humanidades, Andalucía).

Si atendemos a las diferencias según el perfil de estudiante, encontramos que el nivel educativo familiar de los estudiantes tradicionales es superior al de los no tradicionales (ver Gráfico XI).

A menudo, este perfil de estudiantes percibe gran dificultad para afrontar la educación universitaria y tiene carencias académicas de base que arrastran desde años anteriores y les perjudica en la actualidad.

> Bueno, realmente los primeros años me resultó bastante complicado, especialmente el primero porque yo no sabía hacer la O con un canuto en el sentido educativo. Llevaba 15 años sin estudiar y por mucho que llevara 2 años preparándome para entrar a la universidad tenía un nivel muy bajo o al menos así lo sentía yo. (Juan, estudiante no tradicional, ciencias sociales y humanidades, Catalunya).

Recibido: 01/09/2023 Aceptado: 06/06/2024

Además, advierten una dificultad mayor para trabajar en equipo y organizar el estudio (distribuir el tiempo, seguir el ritmo de la clase, etc.).

> Primero sí me resultó complicado, por ejemplo, los trabajos en grupo no sabía cómo hacerlos. Se reunían por el drive y no sabía hacerlo. La manera de buscar información para hacer los trabajos, la bibliografía y todo esto…. No tenía ni idea de cómo hacerlo. La organización para estudiar sí que me costó bastante. (Rocío, estudiante no tradicional, ciencias de la salud, Catalunya).

Experiencia universitaria

Dedicación académica

La dedicación académica de los estudiantes gitanos, al igual que ocurre con estudiantes de minorías étnicas y de primera generación, se ve fuertemente influenciada por la necesidad de equilibrar los estudios con otras responsabilidades, como trabajo y familia (ver Tabla II).

El análisis de los patrones de distribución del tiempo de los participantes refleja particularmente este aspecto entre los estudiantes no tradicionales (ver Gráfico XII). Los estudiantes en general asisten a clase durante aproximadamente 19 horas por semana. Sin embargo, existen

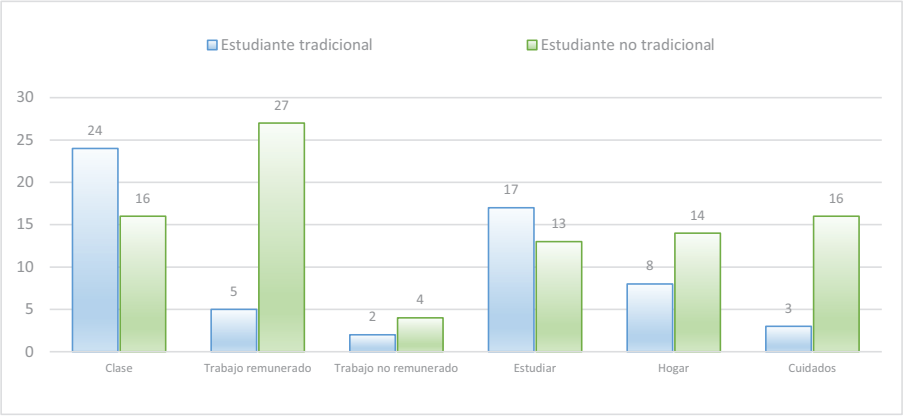
TABLA II. Dedicación en horas a tareas académicas y a otras responsabilidades (%)

	Clase	Trabajo Remunerado	Trabajo Voluntario	Estudiar	Hogar	Cuidados
0h	9%	29%	57%	1%	10%	44%
1-9h	13%	11%	33%	36%	55%	27%
10-19h	28%	18%	9%	34%	16%	5%
20-29h	30%	8%	1%	21%	7%	11%
30-39h	13%	23%	0%	4%	2%	1%
40 o +	6%	11%	0%	4%	9%	11%
MEDIA SEMANA	**19h**	**18h**	**3h**	**15h**	**12h**	**10h**

Fuente: elaboración propia.

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XII. Media a la semana de dedicación a diversas actividades (horas) por perfil de estudiante



Fuente: elaboración propia.

diferencias significativas entre los grupos (24 horas para los estudiantes tradicionales y 16 horas para los no tradicionales). Lo mismo ocurre con las horas dedicadas al estudio y las tareas académicas independientes, con un promedio general de 15 horas por semana, pero una diferencia de 4 horas entre los grupos (17 horas para los tradicionales y 13 horas para los no tradicionales).

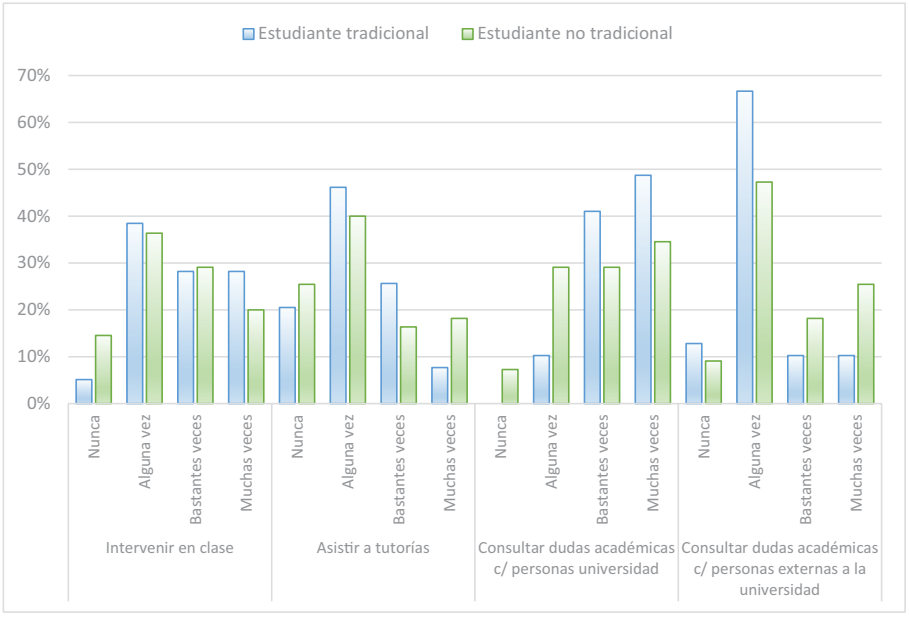
Los datos muestran que este grupo de estudiantes dedica una cantidad significativa de tiempo tanto al trabajo remunerado (18 horas) como a las responsabilidades del hogar (12 horas) y el cuidado de otros (10 horas).

También se manifiesta una diferencia marcada entre los estudiantes tradicionales y no tradicionales en su dedicación a los estudios frente a su dedicación a otras responsabilidades, con promedios de 5 horas, 8 horas y 3 horas para los primeros, frente a 27 horas, 14 horas y 16 horas para los segundos.

Otro de los aspectos analizados en relación con la experiencia universitaria de los estudiantes gitanos muestra que estos tienden a ser participativos en el entorno académico. <t>La gran mayoría de ellos han intervenido en clase y han asistido a tutorías al menos una vez durante el curso en que se obtuvieron los datos. Además, es común que busquen

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XIII. Frecuencia con la que han realizado, en el último curso, diversas actividades, por perfil de estudiante (%)



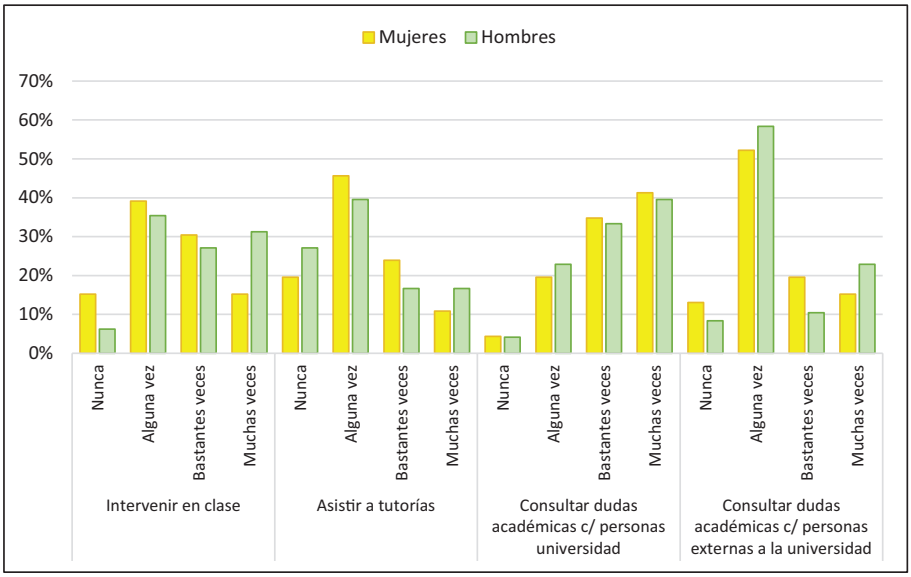
Fuente: elaboración propia.

resolver sus dudas académicas tanto con sus compañeros de clase como con los profesores e incluso con personas externas a la universidad.

La clasificación según los perfiles revela ciertas disparidades en las formas de participación (ver Gráfico XIII). Por ejemplo, los estudiantes tradicionales son más activos en las discusiones en clase, mientras que los no tradicionales tienden a acudir más a tutorías. También observamos que los estudiantes tradicionales suelen plantear sus dudas con mayor frecuencia a personas dentro de la universidad, mientras que los no tradicionales tienden a buscar respuestas mediante la ayuda de personas externas a la institución.

Si bien la distribución por sexos es bastante similar, los resultados muestran diferencias en relación con las intervenciones en clase, teniendo los estudiantes varones mayor tendencia que las mujeres a hacerlo (ver Gráfico XIV).

GRÁFICO XIV. Frecuencia con la que han realizado, en el último curso, diversas actividades, por sexo (%)



Fuente: elaboración propia.

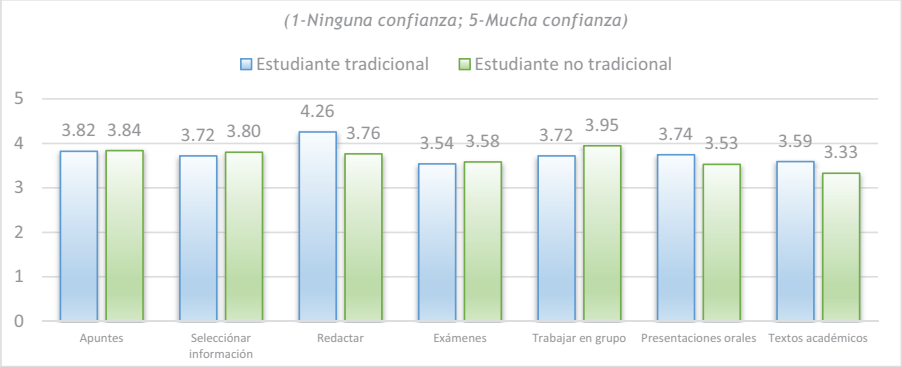
Autoeficacia

La literatura científica documenta que los estudiantes pertenecientes a minorías étnicas, especialmente aquellos con trayectorias no académicas y que ingresaron tardíamente a la universidad, suelen tener una baja confianza en sus habilidades para realizar actividades académicas (Chaney, 2018; Wong, 2018). Sin embargo, los resultados de nuestro estudio parecen contradecir esta percepción, ya que los entrevistados muestran una alta confianza en la mayoría de los aspectos académicos considerados. En todos los casos, las calificaciones medias superan 3 sobre 5 (ver Gráfico XV).

En cuanto al estrés percibido frente a situaciones académicas habituales, como hacer un examen o una presentación oral, los resultados muestran que lo que genera mayor estrés es la realización de exámenes, con un promedio de casi 4 sobre 5. A bastante distancia, encontramos

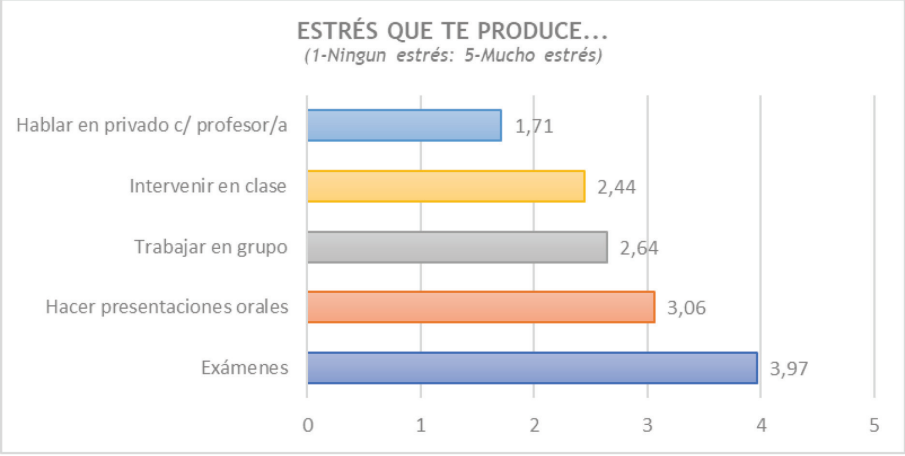
Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XV. Confianza académica por perfil de estudiante



Fuente: elaboración propia.

GRÁFICO XVI. Estrés ante situaciones académicas



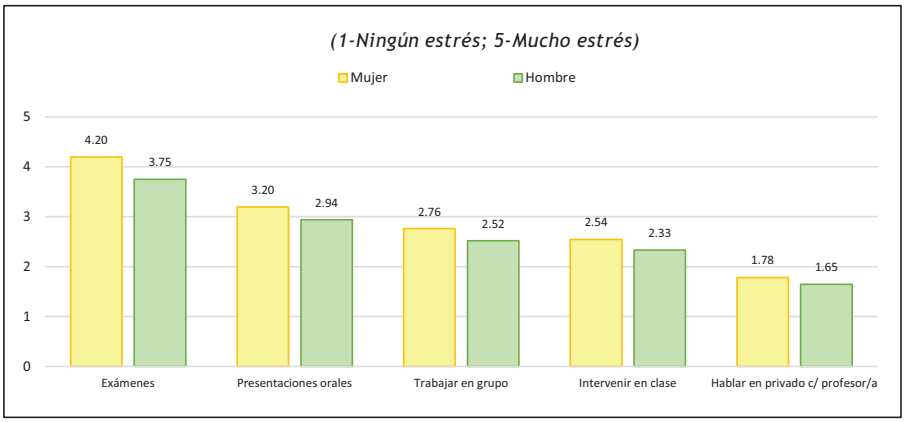
Fuente: elaboración propia.

hacer presentaciones orales, seguida de trabajar en grupo, intervenir en clase y hablar en privado con el profesorado (ver Gráfico XVI).

El análisis por sexo indica que las actividades académicas examinadas generan un grado de estrés significativamente más alto entre las mujeres

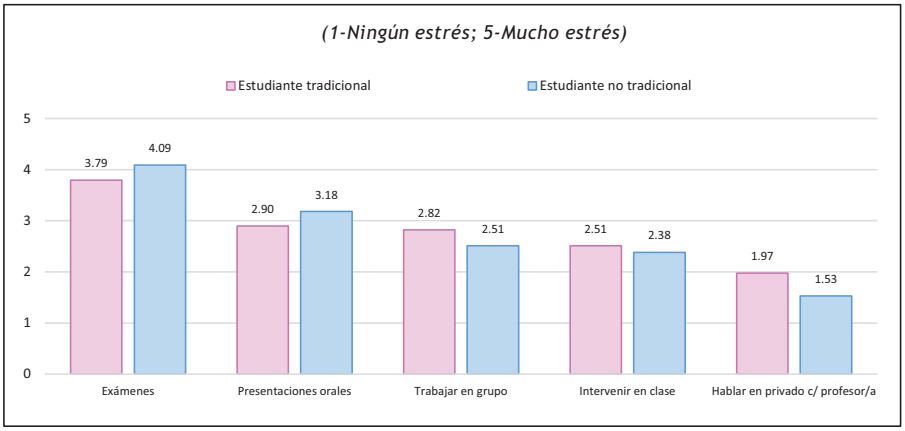
Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XVII. Estrés ante situaciones académicas por sexo



Fuente: elaboración propia.

GRÁFICO XVIII. Estrés ante situaciones académicas por perfil del estudiante



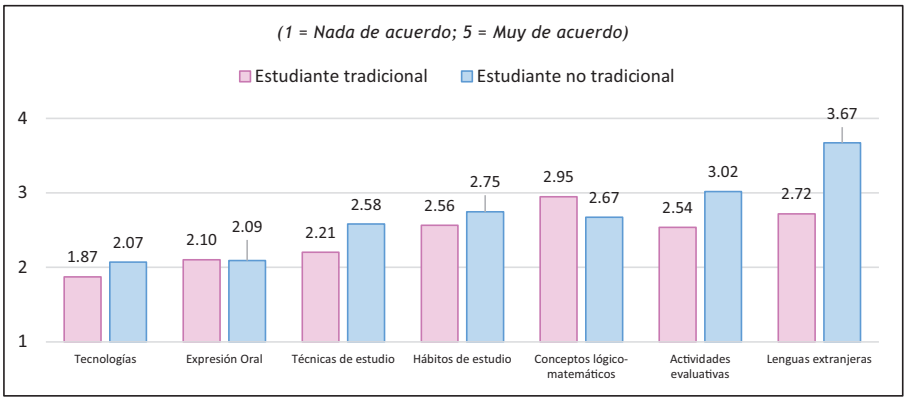
Fuente: elaboración propia.

(ver Gráfico XVII). Sin embargo, los resultados no muestran diferencias reseñables entre ambos perfiles de estudiantes (ver Gráfico XVIII).

Según la percepción de los estudiantes, la competencia académica que representa mayor dificultad sería el dominio de lenguas extranjeras

Recibido: 01/09/2023 Aceptado: 06/06/2024

GRÁFICO XIX. Dificultades académicas por perfil del estudiante



Fuente: elaboración propia.

(ver Gráfico XIX). Este aspecto supone un gran reto especialmente para los estudiantes no tradicionales, teniendo en cuenta que el hecho de no haber seguido trayectorias escolares continuas ha dificultado alcanzar un nivel de inglés adecuado a los requerimientos actuales de la educación superior en el contexto internacional. En relación con la dificultad percibida sobre los conceptos lógico-matemáticos se contempla mayor variabilidad. Por ejemplo, los estudiantes que cursan grados con asignaturas que requieren este tipo de competencias lo señalan como una dificultad importante. Sin embargo, este mismo aspecto es mostrado como carencia entre aquellos estudiantes que cursan grados donde no se requieren competencias lógico-matemáticas. Además, los participantes perciben la necesidad de desarrollar más las técnicas y hábitos de estudio. Y la expresión oral y la tecnología se consideran menos problemáticas.

Conclusiones

El estudio realizado muestra la interseccionalidad de diversos factores de desigualdad que recaen sobre los estudiantes universitarios gitanos.

Específicamente, la discriminación étnica se puede conjugar con la discriminación por razón de sexo y a menudo con otras variables relacionadas con el hecho de ser en su mayoría estudiantes de primera generación, estudiantes adultos y de nivel socioeconómico desfavorable. Pertenecer con mayor frecuencia a un perfil de estudiantes no tradicionales conlleva en muchos casos la necesidad de compaginar los estudios con obligaciones laborales y/o familiares. Los estudiantes perciben que la universidad es excesivamente rígida y está diseñada exclusivamente para estudiantes tradicionales, dando lugar al afrontamiento de grandes dificultades para aquellas personas que se encuentran fuera del modelo tradicional. Paralelamente, el hecho de haber tenido frecuentemente trayectorias educativas discontinuas introduce dificultades específicas relacionadas con lagunas formativas en algunos aprendizajes que han ido ganando importancia en el ámbito de la educación durante los últimos años, como puede ser el dominio de la lengua inglesa y el uso de las tecnologías, o el desarrollo competente de técnicas de estudio.

En relación con la posición socioeconómica de la comunidad gitana en España, encontramos que muchos de ellos disponen de recursos materiales insuficientes. Esto se manifiesta en la dependencia de becas y la necesidad de trabajar para sufragar los estudios. La insuficiencia de recursos para el estudio puede incluir aspectos como la falta de un espacio de trabajo, buena conexión a internet u ordenador de uso propio, aspectos que han sido de mayor relevancia los últimos años, debido al desplazamiento a formatos online o híbridos como consecuencia de las medidas adoptadas para hacer frente a la pandemia.

Históricamente, las estimaciones sobre la población gitana en España se han obtenido mediante trabajos de campo específicos (Ruiz, 2009). De este modo, la escasez de fuentes documentales y la inexistencia de un censo de población oficial dificultan la posibilidad de obtener un conocimiento real de la población gitana en España mediante procedimientos empíricos. No obstante, las limitaciones para generar un marco referencial de muestreo para la población gitana y, concretamente, para identificar estudiantes gitanos universitarios, fueron abordadas siguiendo las indicaciones de la investigación psicosocial para acceder a las poblaciones de difícil acceso (Raifman et al. 2022). Consecuentemente, utilizamos el muestreo de bola de nieve para contactar con estudiantes universitarios gitanos a través de asociaciones y redes sociales romaníes, cumpliendo con el objetivo de identificar las características

Recibido: 01/09/2023 Aceptado: 06/06/2024

sociodemográficas y académicas de los estudiantes universitarios gitanos en una aproximación inicial exploratoria.

La discriminación hacia los grupos étnicos minoritarios es una realidad de la cual no está exenta la universidad. Concretamente, el pueblo gitano es el colectivo que despierta mayor rechazo en nuestra sociedad, por encima de otros colectivos altamente estigmatizados. Como consecuencia, la multiplicidad de barreras de orígenes distintos analizadas en el proyecto UNIROMA son experimentadas por los estudiantes gitanos en mayor proporción que la media, debido a la situación de exclusión social en la que se encuentra la comunidad gitana. Concretamente, los estudiantes gitanos experimentan algún tipo de discriminación a lo largo de su trayectoria universitaria, que, junto a otras dificultades asociadas a la escasez de recursos y la necesidad de compaginar los estudios con otras actividades, puede dificultar la finalización de sus estudios.

El colectivo universitario romaní presenta un perfil no tradicional en una elevada proporción. De este modo, a la discriminación étnica se le añaden otros ejes de desigualdad que operan en el marco universitario, como la edad o la clase social. El riesgo de abandono de los estudios universitarios se incrementa cuando interseccionan las variables de sexo, etnia, estatus socioeconómico, primera generación y perfil no tradicional. Por lo tanto, los resultados muestran la necesidad de realizar acciones que garanticen la permanencia del alumnado romaní, en su gran mayoría de primera generación, para finalizar los estudios superiores. El proyecto UNIROMA ha formulado recomendaciones basadas en la evidencia para contribuir a este objetivo. Sin embargo, se requieren futuras investigaciones que sigan aportando conocimiento para mejorar la situación del alumnado gitano y las barreras institucionales a las que se enfrentan en la educación superior.

Este artículo añade información relevante y significativa sobre la poca representación de estudiantes de primera generación y/o pertenecientes a minorías étnicas en las universidades españolas y específicamente sobre las tasas de graduación, aspectos que preocupan también a nivel internacional. Para graduarse, este alumnado requiere de un apoyo integral a través de mentorías, apoyo académico, económico, orientación y fomento de su participación y vínculo con la universidad, para poder concluir sus estudios.

La falta de representación de los estudiantes romaníes en la educación superior se explica por una desigualdad histórica y estructural que

requiere de una responsabilidad social por parte de los gobiernos y las instituciones de educación superior. Aspectos a tener en cuenta son la falta de familiaridad con el entorno universitario y los efectos demoledores de carecer de un sentido de pertenencia. Los universitarios y titulados superiores de minorías y de primeras generaciones resultan modelos a seguir para el resto de la comunidad, pues inspiran y animan a futuros estudiantes (Aiello et al., 2019). Incorporar sus voces para mejorar las políticas permite identificar mejor las dificultades que experimentan, pero también ayuda a que la información llegue a las nuevas generaciones de universitarios. Así, los primeros gitanos en finalizar una titulación universitaria constituyen referentes muy importantes para su pueblo y es un aspecto que las acciones a realizar en las universidades deberán tener en cuenta.

La escasez de conocimiento e información sobre la situación del pueblo gitano en España y Europa, tanto del profesorado como del personal de administración y servicios, dificulta el acompañamiento de los estudiantes gitanos. Acciones formativas que respondan a esta carencia y mejoren la representación de la comunidad gitana en la universidad entre el profesorado pueden contribuir a que los estudiantes gitanos conecten con la institución y creen un vínculo.

Agradecimientos

Agradecemos a todos los miembros del equipo investigador (Silvia Abad Merino, Jelen Amador, Ainhoa Flecha, Tania García, Sonia Garcia Segura, Ana Giménez, Fernando Macías, Paz Peña, Blas Segovia, Marta Soler, Teresa Sordé.) así como a los miembros del Consejo Asesor (Federación KAMIRA, CAMPUSROM, Santiago Guerrero, Antonio Madrid, Itxaso Tellado, Ramon Vílchez, Rafael Santiago, Aaron Giménez, Jesús Martínez, Loli Santiago, Sara Santiago) y a todos los y las estudiantes que han contribuido con su experiencia en este proyecto.

Financiación

Este estudio ha sido financiado por el Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020 del Ministerio de

Ciencia, Investigación y Universidades, con cofinanciación de los Fondos Estructurales de la Unión Europea. (Referencia RTI2018-097879-A-100).

Referencias bibliográficas

Agencia de los Derechos Fundamentales de la Unión Europea (EU-FRA). (2016). *Second European Union Minorities and Discrimination Survey (EU-MIDIS II) Roma – Selected Findings*. Vienna: European Union Agency for Fundamental Rights. https://doi.org/10.2811/469.

Agencia de los Derechos Fundamentales de la Unión Europea (EU-FRA). (2022). *Roma in 10 European Countries*. Recuperado de https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-roma-survey-2021-main-results2_en.pdf.

Aiello, E.; Amador-López, J.; Munté-Pascual, A.; & Sordé-Martí, T. (2019). Grassroots Roma Women Organizing for Social Change: A Study of the Impact of 'Roma Women Student Gatherings'. *Sustainability*, *11*, 4054. https://doi.org/10.3390/su11154054

Alexiadou, N. (2019). Framing education policies and transitions of Roma students in Europe. *Comparative Education*, *55*(3), 422-442. https://doi.org/10.1080/03050068.2019.1619334.

Allaire, F. S. (2019). Navigating uncharted waters: First-generation native Hawaiian college students in STEM. *Journal of College Student Retention: Research, Theory and Practice*, *21*(3), 305-325. https://doi.org/10.1177/1521025117707955.

Amnesty International y European Roma Rights Center. (2017). *A Lesson in Discrimination: Segregation of Romani Children in Primary Education in Slovakia*. Recuperado de http://www.errc.org/uploads/upload_en/file/report-lesson-in-discrimination-english.pdf.

Arday, J., & Mirza, H. S. (Eds.). (2018). *Dismantling race in higher education: Racism, whiteness and decolonizing the academy*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-60261-5.

Bîrle, D., Crişan, D., Bonchiş, E., Bochiş, L., & Popa, C. (2014). Roma Social Inclusion through Higher Education Policies in Romania. En F. Ribeiro, Y. Politis y B. Culum (Eds.), *New Voices in Higher Education Research and Scholarship* (pp. 213-231). IGI Global. https://doi.org/10.4018/978-1-4666-7244-4.ch011.

Brüggemann, C. (2014). Romani culture and academic success: arguments against the belief in a contradiction. *Intercultural Education, 25*(6), 439-452. https://doi.org/10.1080/14675986.2014.990229.

Brüggemann, C., & D'Arcy, K. (2017). Contexts that discriminate: International perspectives on the education of roma students. *Race Ethnicity and Education, 20*(5), 575-578. https://doi.org/10.1080/13613324.2016.1191741.

Cashman, L. (2017). New label no progress: Institutional racism and the persistent segregation of Romani students in the Czech Republic. *Race Ethnicity and Education, 20*(5), 595-608. https://doi.org/10.1080/13613324.2016.1191698.

Center for Postsecondary Research (2019). NSSE (National Survey of Students' Engagement) Modules. Recuperado de: https://nsse.indiana.edu/html/topical_module_participation.cfm [Consultado el 18 de marzo de 2019].

Chaney, I. (2018). Confidence in Ability: The Non-Traditional University Student. In *11th International Conference of Education, Research and Innovation, Seville, 12-14 November 2018* (pp. 7206-7211).

Comisión Europea. (2010). *Europe 2020: A European strategy for smart, sustainable and inclusive growth.*

Comisión Europea. (2020). *A union of Equity: EU Roma strategic framework for equality, inclusion and participation for 2020-2030.*

CREA. (2010). *Gitanos: de los mercadillos a la escuela y del instituto al futuro.* Madrid: Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación.

De la Rica, S., Gorjón, L., Miller, L., & Úbeda, P. (2019). *Estudio comparado sobre la situación de la población gitana en España en relación al empleo y la pobreza*. FSG. Recuperado de https://iseak.eu/wp-content/uploads/2019/09/Estudio-comparado-sobre-la-situación-de-la-población-gitana-en-España-en-relación-al-empleo-y-la-pobreza.pdf.

Flecha, R. (2018). Comunidades de aprendizaje y transformación social. *Educadores: revista de renovación pedagógica*, *265*, 44-54.

Flecha, A., Abad-Merino, S., Macías-Aranda, F., & Segovia-Aguilar, B. (2022). Roma University Students in Spain: Who Are They? *Education Sciences, 12*(6), 400. https://doi.org/10.3390/ educsci12060400

Fundación Europea para la mejora de las condiciones de vida y de trabajo (Eurofound). (2016). *Exploring the Diversity of NEETs.* Publications Office of the European Union. https://doi.org/10.2806/62307.

Recibido: 01/09/2023 Aceptado: 06/06/2024

Garaz, S., & S. Torotcoi (2017). Increasing access to higher education and the reproduction of social inequalities: The case of Roma university students in eastern and southeastern Europe. *European Education*, *49*(1), 10-35. https://doi.org/10.1080/10564934.2017.1280334.

Gargallo, B., Jesús, S. R., & Pérez-Pérez, C. (2009). El cuestionario CEVEAPEU. Un instrumento para la evaluación de las estrategias de aprendizaje de los estudiantes universitarios. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, *15*(2), 1-31.

Gómez, A., Racionero, S. & Sordé, T. (2010). Ten years of critical communicative methodology. *International Review of Qualitative Research, 3*(1), 17-43.

Harris, J. C., & Linder, C. (2018). The racialized experiences of students of color in higher education and student affairs graduate preparation programs. *Journal of College Student Development*, *59*(2), 141-158. https://doi.org/10.1353/csd.2018.0014.

Hinton-Smith, T., Danvers, E., & Jovanovic, T. (2017). Roma women's higher education participation: whose responsibility? *Gender and Education*, *38*(7), 811–828. https://doi.org/10.1080/09540253.2016.1274386.

Hinton-Smith, T., & Padilla-Carmona, M. T. (2021). Roma university students in Spain and Central and Eastern Europe: Exploring participation and identity in contrasting international contexts. *European Journal of Education*, *56*(3), 454-467.https://doi.org/10.1111/ejed.12459.

Kamberi, E., Martinovic, B., & Verkuyten, M. (2015). Life satisfaction and happiness among the Roma in Central and southeastern Europe. *Social Indicators Research*, *124*(1), 199-220.

Kende, A., Hadarics, M., Bigazzi, S., Boza, M., Kunst, J. R., Lantos, N. A., Lášticová, B., Minescu, A., Pivetti, M., & Urbiola, A. (2020). The last acceptable prejudice in Europe? Anti-Gypsyism as the obstacle to Roma inclusion. *Group Processes & Intergroup Relations*. https://doi.org/10.1177/1368430220907701.

Kundu, A. (2019). Understanding college "Burnout" from a social perspective: Reigniting the agency of low-income racial minority strivers towards achievement. *Urban Review*, *51*, 667-698. https://doi.org/10.1007/s11256-019-00501-w.

Lambrev, V. (2020). 'They clip our wings': studying achievement and racialisation through a Roma perspective. *Intercultural Education*, *31*(2), 139-156. https://doi.org/10.1080/14675986.2019.1702292.

Lens, D., & Levrau, F. (2020). Can pre-entry characteristics account for the ethnic attainment gap? An analysis of a Flemish university. *Research in Higher Education*, *61*(1), 26-50.https://doi.org/10.1007/S11162-019-09554-Y.

Marcu, S. (2020). Mobility as a learning tool: educational experiences among Eastern European Roma undergraduates in the European Union, *Race Ethnicity and Education*, *23*(6), 858-874. https://doi.org/10.1080/13613324.2019.1579179.

Márquez, J.L., & Ordaz, M. (2014). *Caracterización psicopedagógica del estudiante universitario. Cuestionario para el ingreso*. Congreso Universidad. Vol. III, No. 2, 1-14, ISSN: 2306-918X.

McDonagh, C., & Fonseca, J. (2022). The minority within the minority: The survival strategies of Gypsy and Traveler students in Higher Education. *International Journal of Roma Studies, 4*(2), 150-171. http://dx.doi.org/10.17583/ijrs.10344.

Messing, V. (2017). Differentiation in the making: Consequences of school segregation of Roma in the Czech Republic, Hungary, and Slovakia. *European Education, 49*(1), 89-103. https://doi.org/10.1080/1056493 4.2017.1280336.

Ministerio de Derechos Sociales y Agenda 2030. (2021). *Estrategia nacional para la Igualdad, Inclusión y Participación de la Población Gitana en España 2021-2030*. Recuperado de https://www.mdsocialesa2030.gob.es/derechos-sociales/poblacion-gitana/docs/estrategia_nacional/Estrategia_nacional_21_30/estrategia_aprobada_com. pdf.

Mulcahy, E., Baars, S., Bowen-Viner, K., & Menzies, L. (2017). *The underrepresentation of Gypsy, Roma and Traveler pupils in higher education. A report on barriers from early years to secondary and beyond. King's College London*. Recuperado de https://www.lkmco.org/wp-content/uploads/2017/07/KINGWIDE_28494_proof3.pdf.

National Center for Education Statistics (NCES). (2024). *Nontraditional Undergraduates/Highlights*. Recuperado de: https://nces.ed.gov/pubs/web/97578a.asp.

Neumann, E. (2016). 'Fast and violent integration': school desegregation in a Hungarian town, *Race Ethnicity and Education*, *20*(5), 579-594. https://doi.org/10.1080/13613324.2016.1191696.

OCDE (2020). *Population with tertiary education* (indicator). https://doi.org/10.1787/0b8f90e9

Ogunyemi, D., Clare, C., Astudillo, Y. M., Marseille, M., Manu, E. y Kim, S. (2019). Microaggressions in the learning environment: A systematic

review. *Journal of Diversity in Higher Education*, *13*(2), 97-119. https://doi.org/10.1037/dhe0000107.

Parra Toro, I., Álvarez-Roldán, A., & Gamella, J. F. (2017). A Silenced Conflict: Segregation Processes, Curricular Gaps and Dropout among Spanish Gitanos or Cale. *Revista de paz y conflictos*, *10*(1), 35-60. https://doi.org/10.30827/revpaz.v10i1.5965.

Raifman, S., DeVost, M.A., Digitale, J.C. et al. (2022). Respondent-Driven Sampling: a Sampling Method for Hard-to-Reach Populations and Beyond. *Curr Epidemiol Rep, 9*, 38–47. https://doi.org/10.1007/s40471-022-00287-8

Richardson, J. T. E. (2018). Understanding the under-attainment of ethnic minority students in UK higher education: The known knowns and the known unknowns. En *Dismantling race in higher education: Racism, whiteness and decolonising the academy* (pp. 87-102). https://doi.org/10.1007/978-3-319-60261-5_5.

Roberts, P. (2021). Class dismissed: international mobility, doctoral researchers, and (Roma) ethnicity as a proxy for social class? *Discourse: Studies in the Cultural Politics of Education*, *42*(1), 142-154. https://doi.org/10.1080/01596306.2020.1855569.

Ruiz, J. S. (2009). Orígenes, vicisitudes, realidad actual y retos del pueblo gitano en España y Región de Murcia. In *Anales de Historia Contemporánea*, *25*, 115-131.

Rutigliano, A. (2020). Inclusion of Roma students in Europe: A literature review and examples of policy initiatives.*OECD Education Working Paper No. 228***.** Recuperado de https://www.oecd-ilibrary.org/education/inclusion-of-roma-students-in-europe_8ce7d6eb-en

Salamońska, J. y Czeranowska, O. (2018). *How to research multiple migrants? Introducing web-based respondent-driven sampling survey.* CMR Working Papers, No. 110/168, University of Warsaw, Centre of Migration Research (CMR), Warsaw.

Schanze J.-L., & Zins S. (2019). Undercoverage of the elderly institutionalized population: The risk of biased estimates and the potentials of weighting. *Survey Methods: Insights from the Field*. Recuperado de: https://surveyinsights.org/?p=11275 [Consultado del 11 de marzo de 2019]

Serradell, O., García-Cabrera, M. M., Barrera-Rodríguez, M. F., & Ramírez, C. (2022). Recomendaciones para el apoyo de estudiantes Romanís en la Universidad. *International Journal of Roma Studies, 4(*1), 172- 203. https://doi.org/10.17583/ijrs.10682.

Silver, B. R. (2019). On the margins of college life: The experiences of racial and ethnic minority men in the extracurriculum. *Journal of Contemporary Ethnography*, *49*(2), 147-175. https://doi.org/10.1177/0891241619869808.

Slobodnikova, A, & Randolph-Seng, B. (2021). The effects of stereotype threat on Roma academic performance in Slovakia: role of academic self-efficacy and social identity. *Journal for Multicultural Education*, *15*(2), 152-167. https://doi.org/10.1108/JME-08-2020-0080.

Students for Education Reform (2017). *First Generation College Survey.* Recuperado de: https://d3n8a8pro7vhmx.cloudfront.net/sfer/pages/5254/attachments/original/1494530447/FGCS_Poll_5-11-17.pdf?1494530447 [Consultado el 18 de marzo de 2019]

The Velux Foundations (2019). "Education", in *2019: Roma in Europe*.

Times Higher Education – World University Rankings (2018). *Student Experience Survey – Methodology.* Recuperado de: https://www.timeshighereducation.com/student/news/student-experience-survey-2018-methodology [Consultado el 18 de marzo de 2019]

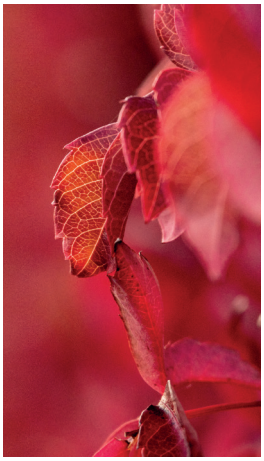
Urbiola, A., Navas, M., Carmona, C. et al. (2022). Social Class also Matters: The Effects of Social Class, Ethnicity, and their Interaction on Prejudice and Discrimination Toward Roma. *Race and Social Problems*, 1-15. https://doi.org/10.1007/s12552-022-09368-1.

Wong, B. (2018). By Chance or by Plan? The Academic Success of Nontraditional Students in Higher Education. *AERA Open*, *4*(2). https://doi.org/10.1177/2332858418782195

Želinský, T. (2022). Stereotype threat among European Roma adults. *Applied Economics Letters*, *29*(1), 49-52, https://doi.org/10.1080/13504851.2020.1855301

Información de contacto: Fernando Macías Aranda. Passeig de la Vall d'Hebrón, 171. Despacho 217. 2a planta. Edifici Llevant. 08035 Barcelona. E-mail: fernandomacias@ub.edu

Recibido: 01/09/2023 Aceptado: 06/06/2024



Reseñas

Thoilliez, B. y Manso, J. (Coords.) (2023). *La Educación, ¿bien común, derecho individual o servicio público?* Madrid: Editorial Sintesis. 147 pp. ISBN:978-84-1357-265-9

Actualmente, un debate muy extendido en el ámbito educativo es la concepción de lo público. A pesar de que en numerosas ocasiones este se centre únicamente en la gestión económica de los centros, existe una dimensión filosófica y teórica más profunda. El libro *La Educación, ¿bien común, derecho individual o servicio público?* aborda fundamentos teóricos que nos aclaran los nuevos horizontes para repensar la escuela como institución que extienda y asegure los derechos y las oportunidades del educando. En otras palabras, la escuela como un bien público y común.

A lo largo del primer capítulo, Miriam Prieto y Alberto Sánchez ahondan en las cuestiones relacionadas con el bien y lo común. Añaden la importancia de expresar el reconocimiento de las minorías sin caer en el relativismo, ni privatizando esta idea de alcanzar el bien común.

En el segundo capítulo, Fernando Gil Cantero defiende una perspectiva pedagógica de lo político frente a la perspectiva política de lo pedagógico. En la primera prima el bien, sin embargo, en la segunda perspectiva prevalecen las negociaciones y, por tanto, intereses. Explora distintos aspectos de la educación desmitificando la dualidad generada por la polarización ideológica de la Pedagogía.

David Reyero es el encargado del tercer capítulo, donde a través de la ética y antropología liviana reconoce concepciones relacionadas con la búsqueda del *telos*. El autor no pierde de vista el espacio educativo mediante el pluralismo y la densidad para alcanzar la virtud, el bien y la verdad. También analiza concepciones de la democracia y ontológicas que han surgido en la Modernidad. Este subjetivismo egocéntrico que ha ido surgiendo mediante esa exaltación de la individualidad ha generado unas concepciones de lo diferente. El entramado ético-político se puede tratar desde aportaciones de Taylor pues nos aproximan a lo deseable. En el cuarto capítulo Tania Alonso trata dichas aportaciones y las relaciona con la promoción de la escuela en lo bueno y éticamente superior, como un proyecto compartido y un destino común.

El sistema neoliberal ha introducido una cultura pop centrada en la felicidad y la diversidad. A lo largo del quinto capítulo Bianca Thoilliez explora la repercusión de este ensimismamiento de las modas educativas porque privatizan los bienes escolares. La autora reivindica la importancia

de apropiarnos de la cultura común y diferencia la profesión docente frente al terapeuta (siendo la del primero un ejercicio continuado de redistribución).

Los distintos proyectos que surgen de la nueva educación para satisfacer a un colectivo de familias (buscan la felicidad y rechazan la directividad bajo la premisa de *laissez-faire*) son, como expresa Ani Pérez en el capítulo seis, una amenaza para la construcción de una escuela hacia el bien común. A lo largo de este capítulo nos aproximamos desde otra perspectiva a lo que supone la libertad de centro en un contexto de mercado. Este contexto de mercantilización y privatización se aborda en los dos siguientes capítulos; en el séptimo Mariano Narodowski y Delfina Campetella tratan el escenario de *razón de Estado* y las barreras que la conducen hacia una *razón de mercado*, por último, Jesús Manso y Marta Moreno identifican los principios de la Nueva Gestión de lo Público introducidos como una *soft policy* especialmente analizados en la normativa educativa española de los últimos años.

La obra culmina con un epílogo titulado *la escuela, un bien público y común*, donde Esther Díaz Romanillos recopila los conceptos trabajados como; (i) lo público y lo privado, lo común y el bien; (ii) las políticas de la identidad; (iii) la libre elección de centros y las relaciones de poder. Este libro aporta y eleva el debate sin perder de vista lo que realmente importa: la escuela "como un espacio educativo democrático único, donde la democracia se concreta en una actividad y movimiento singular y liberador" (p. 147).

Si queremos lograr una sociedad más justa debemos defender la escuela y, esto se hace sin culparla de todos los males de la sociedad, pues las demás instituciones también se deben responsabilizar para que la ciudadanía alcance lo deseable para una *vida buena*.

Carmen Fontaneda Amo

Bautista García-Vera, A. (2021). Audiovisuales, desigualdades socioculturales y educación. *Educatio Siglo XXI*, *41*(1), 149-154. Ediciones de la Universidad de Murcia. ISBN: 978-84-18936-27-2

"Audiovisuales, desigualdades socioculturales y educación" es un ensayo basado en una investigación de paradigma interpretativo que muestra cómo las tecnologías audiovisuales y la brecha digital influencian en el aumento de las desigualdades socioculturales en el sistema educativo. Antonio Bautista García-Vera, catedrático de Didáctica y Organización Escolar, expresa entre líneas su sensibilidad y preocupación por la mejora de la educación escolar a través de las posibilidades de las tecnologías audiovisuales. El libro sigue un discurso y estructura propios de un trabajo de investigación, distribuyendo el contenido en ocho capítulos, anexos y bibliografía.

El autor nos adentra en la investigación con la narración en primera persona de la toma de contacto inicial con los centros educativos. En el primer capítulo, el lector se sumerge en la cotidianidad de los dos centros de educación primaria de la Comunidad de Madrid, protagonistas del que iba a devenir un nuevo escenario para la generación de conocimiento compartido. Se cuenta con la colaboración entre la comunidad educativa –profesorado, madres, padres y alumnado- y el equipo investigador.

El segundo y tercer capítulo del libro se destinan a guiar al lector en la comprensión de los propósitos concretos de la investigación. Los antecedentes señalan, primeramente, la importancia de diferenciar aquellos usos más técnicos y funcionales que presenta una herramienta digital, del papel o función que adquiere en la sociedad y en la educación o de su significado. Los marcos conceptuales interrelacionados guían a comprender cómo el significado que les otorgamos a las tecnologías se vincula en las desigualdades socioculturales y educativas. Se citan autores reconocidos e investigaciones emergentes en referencia a brecha digital, la alfabetización tecnológica, el arte en la sensibilización estética y la competencia digital.

Desde este estado del arte, surge el origen del estudio bajo tres curiosidades a resolver acerca de las herramientas tecnológicas y audiovisuales en la educación formal: comprender el significado que adquieren en los centros educativos; conocer cómo se favorece el desarrollo personal, social y académico del alumnado; e identificar las implicaciones en el desarrollo profesional y personal docente. Para

responder a ellos, el cuarto capítulo se destina a describir los principios de recogida de datos, el análisis del discurso y las fases del estudio. Los métodos de la investigación se concretan de forma minuciosa y detallada, considerando aspectos éticos, de coherencia y de validez. La recogida de datos mediante entrevistas se dio durante y después de la implementación de procesos de alfabetización digital para profesorado y alumnado. El estudio denota rigurosidad y pertinencia metodológica en la forma en que se interpretan y se describen las sensaciones, percepciones, análisis y reflexiones del alumnado y profesorado.

Las tres curiosidades se resuelven en los capítulos 5, 6 y 7, respectivamente. Los resultados demuestran cómo las funciones que se asignaban a los medios se habían ampliado durante el estudio: de significados sobre la transmisión de información y actividades de diversión, hacia significados vinculados con la expresión, la creatividad y la resolución de problemas. Este es un gran hallazgo del proyecto, ya que se muestra cómo promover el logro de las Competencias Digitales en contextos educativos multiculturales para la inclusión social. Otro hallazgo significativo es el impacto que ha generado la alfabetización audiovisual en el rendimiento académico, el desarrollo personal y social, y la proyección profesional del alumnado. Las producciones audiovisuales han causado una mejora en la sensibilidad por lo bello y el arte en vinculación a experiencias cotidianas y propios intereses. Los resultados también reflejan el rol de los medios en el desarrollo personal y el desarrollo profesional del profesorado. Éstos adquieren protagonismo cuando el profesorado trata de conectar con los intereses y motivaciones del alumnado. Para lectores del libro de Bautista, resulta de especial interés detenerse a revisar en detalle las citas de las entrevistas de los participantes.

La esencia de los resultados se presenta en el octavo y último capítulo en forma de conclusiones. Bautista reflexiona con profundidad acerca de las contribuciones del estudio para la disminución de la brecha digital y la mejora de vínculos sociales a través el desarrollo tecnológico en comunidades educativas. La transparencia en la descripción de la investigación, la sensibilidad por el mundo audiovisual y la preocupación por la mejora de la educación guían los motivos por adentrarse en este libro.

Laura Fernández-Rodrigo

Moliner, O. (2024). *Democratizar la universidad. Retos de la educación inclusiva en la formación postsecundaria de jóvenes con discapacidad intelectual*. Barcelona: Octaedro. 246 pp. ISBN: 9788419900999

Soñar con una universidad inclusiva que reconozca, dignifique y vele por los derechos de las personas con discapacidad no es una utopía, sino una cuestión de humanidad, responsabilidad y justicia social. Esta obra coordinada por Odet Moliner está elaborada desde la coherencia y el compromiso firme con la inclusión desde voces polifónicas, presentando no solo análisis críticos sobre este desafío comunitario, sino también propuestas para la acción desde las universidades.

Este libro narra los resultados de un proyecto de construcción colectiva del conocimiento por parte de estudiantes con discapacidad intelectual (DI), representantes del tejido asociativo y la Administración, docentes, profesionales e investigadores universitarios. Se trata de un proyecto enmarcado en el Programa UniDiversitat financiado por la Fundación ONCE y el Fondo Social Europeo, que se oferta como Título Propio de educación postsecundaria en la Universitat Jaume I para promover la vida independiente, así como la inclusión social y laboral de jóvenes con DI inscritos en el Sistema de Garantía Juvenil.

La estructura de esta obra está compuesta por diez capítulos y consta de dos partes diferenciadas. La primera titulada "Recorridos" contiene seis capítulos y la segunda enunciada "Tránsitos" engloba cuatro capítulos. El primer capítulo sienta las bases sobre cómo la universidad puede ser inclusiva democratizando tres funciones: formativas (pedagogía crítica, transformadora y abierta al cuidado), investigadoras (investigación inclusiva, participativa y vinculada al territorio) y extensionistas (vínculo entre la universidad y la sociedad). El segundo capítulo está desarrollado por investigadores internacionales reconocidos en el contexto irlandés y presenta distintas oportunidades para luchar por las oportunidades de jóvenes con DI en la Educación Superior, como el programa de Artes, Ciencias y Prácticas Aplicadas Inclusivas (ASIAP) y el Foro Nacional de Educación Superior Inclusiva (INHED). El tercer capítulo documenta un proceso de diagnóstico social participativo que identificó, a partir de la cartografía social y técnicas de aprendizaje cooperativo, las necesidades formativas de los jóvenes con DI del programa UniDiversitat de la Universitat Jaume I, promoviendo la participación crítica del entorno a través de la investigación-acción participativa, analizando las

posibilidades de formación o empleo y buscando soluciones conjuntas para ajustar el programa formativo a la realidad social y laboral de los jóvenes. Los capítulos cuarto, quinto y sexto están redactados por el propio el profesorado de la formación que impartía clases inclusivas con estudiantes universitarios y estudiantes con DI del programa UniDiversitat. En estos tres capítulos, se puede apreciar cómo los profesionales diseñaron, desarrollaron y evaluaron propuestas accesibles con futuros docentes y jóvenes con DI sobre microrrelatos, arte y creatividad con talleres musicales, deporte y alimentación saludable.

El séptimo capítulo aporta valor y congruencia a la obra siguiendo los principios de la investigación inclusiva porque está escrito en lectura fácil, con apoyos visuales y tres autores son jóvenes con DI del programa que reflejan sus propios testimonios sobre la empleabilidad. El octavo capítulo realiza una aproximación teórica y práctica sobre el empleo con apoyo, continuando con un workshop comunitario en el marco del programa UniDiversitat acerca de las barreras, ayudas y mejoras de las prácticas laborales. El noveno capítulo presenta la Guía GAS-VI y está elaborado por dos autoras de la Universitat de Girona con una trayectoria reconocida en la vida independiente desde el modelo de derechos de las personas con DI. El décimo capítulo plantea retos futuros, narrando el desarrollo, el empoderamiento, la valoración y las cuestiones aprendidas de un curso de formación en investigación inclusiva "Investiguem" con jóvenes con DI del programa UniDiversitat sobre temáticas que realmente deciden e implican a las personas con DI en todas sus fases. En definitiva, esta obra debería ser leída detenidamente por toda la ciudadanía y, en particular, por las personas con discapacidad, los líderes de las universidades, el personal técnico de los servicios y las oficinas de discapacidad de las universidades, las empresas, las entidades que trabajan con personas con discapacidad y los familiares de personas con DI. Sin duda, es un libro que recomiendo encarecidamente porque (nos) brinda un halo de esperanza, inspiración y denuncia social, animando a la comunidad universitaria a seguir repensando sus funciones para avanzar en la promoción de la inclusión y movilizar acciones situadas que garanticen la igualdad de oportunidades para todas las personas.

Agradecimientos a la Ayuda Juan de la Cierva-Formación JDC2022-049600-I financiada por MCIU/AEI/10.13039750110001103 3 y por la Unión Europea "Next GenerationEU" /PRTR.

Inmaculada Orozco Almario

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.

GOBIERNO DE ESPAÑA
MINISTERIO DE EDUCACIÓN, FORMACIÓN PROFESIONAL Y DEPORTES

NIPO pdf: 164-24-063-9
NIPO ibd: 164-24-061-8
NIPO html: 164-24-062-3
ISSN línea: 1988-592X
ISSN papel: 0034-8082

https://revistaeducación.es/