

Frade-Martínez, C., Gamazo, A., y Olmos-Migueláñez, S. (2024). Evolución de los predictores contextuales del nivel competencial de las y los estudiantes españoles: un estudio comparativo entre PISA 2015 y 2018. *Revista de Investigación Educativa*, 42(2).  
DOI: <https://doi.org/10.6018/rie.548151>

## **Evolución de los predictores contextuales del nivel competencial de las y los estudiantes españoles: un estudio comparativo entre PISA 2015 y 2018**

### **Evolution of the Contextual Factors Predicting Spanish Students' Competence Levels: A Comparative Study between PISA 2015 and 2018**

Cristina Frade-Martínez<sup>\*1</sup>, Adriana Gamazo\* y Susana Olmos-Migueláñez\*

\*Instituto Universitario de Ciencias de la Educación.  
Universidad de Salamanca (España)

#### **Resumen**

*El Programa para la Evaluación Internacional de Alumnos (PISA) lleva más de 20 años evaluando el nivel competencial del alumnado de 15 años, influyendo a su vez en el establecimiento de políticas y prácticas educativas basadas en sus resultados a nivel internacional. Aunque su configuración no permite el establecimiento de estudios longitudinales, este artículo plantea el diseño de un estudio de tendencias que posibilita la valoración de la evolución de aquellos factores de carácter sociodemográfico o de contexto educativo que mejor predicen el nivel competencial de los y las estudiantes. A través de un análisis de regresión multinivel (modelos jerárquicos lineales) con la muestra española de las ediciones 2015 y 2018 de PISA, conformada por 65684 estudiantes y 1873 centros educativos, se observan los cambios en las variables predictoras de carácter contextual del rendimiento en comprensión lectora, ciencias y matemáticas. Los hallazgos más reseñables son la reducción del impacto del estatus migratorio de los y las inmigrantes de primera generación, la disminución de la brecha de género en las materias STEM (y su aumento en la comprensión lectora) o la reducción del efecto contextual del nivel socioeconómico medio del estudiantado de un centro. Se concluye con la necesidad*

---

<sup>1</sup> **Correspondencia:** Cristina Frade-Martínez, [cristina.frade@usal.es](mailto:cristina.frade@usal.es), Paseo Canalejas 169, 37008, Salamanca, España.

*de realizar análisis más profundos, tanto a nivel estadístico como de política educativa, para poder producir resultados más detallados que permitan esclarecer qué medidas son útiles para la reducción del impacto de los factores socioeconómicos, demográficos y de contexto educativo en el rendimiento de las y los estudiantes españoles.*

*Palabras clave:* PISA; rendimiento; evaluación educativa; educación secundaria.

### **Abstract**

*The Programme for International Student Assessment (PISA) has been assessing student competence levels for over 20 years, while also influencing the implementation of educational policies and practices based on its results at an international level. Although PISA's configuration does not allow for longitudinal studies, this paper proposes the design of a trend study that enables the assessment of the evolution of the sociodemographic and educational context factors that best predict student competence levels. Through a multilevel regression analysis (hierarchical linear models) with the Spanish sample from PISA 2015 and 2018 waves, consisting of 65684 students and 1873 schools, the changes in variables predicting student performance in reading, math and science can be observed.*

*The most noteworthy findings are the reduction of the impact of migration status for first generation immigrant students, the narrowing of the gender gap in STEM subjects (and its widening in reading), or the decrease of the contextual effect of the average socioeconomic level of students in a school. The paper concludes with the need to perform deeper analyses, both at statistical and educational policy levels, to produce more detailed results that shed light on which measures are more useful in order to reduce the impact of socioeconomic, demographic and educational context factors on Spanish students' performance.*

*Keywords:* PISA; performance; educational assessment; secondary education.

### **Introducción y objetivos**

El Programa para la Evaluación Internacional de Alumnos (PISA), desarrollado por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), es una de las pruebas internacionales de evaluación a gran escala que mayor impacto y repercusión ha tenido en las sociedades y en los sistemas educativos desde que se comenzó a implantar en el año 2000.

Este programa se concibió como un “recurso para ofrecer información abundante y detallada que permita a los países miembros adoptar las decisiones y políticas públicas necesarias para mejorar los niveles educativos” (OECD, 2006, p. 3). Ciertamente, desde su primera aplicación, los resultados de PISA han generado diversas reacciones en los gobiernos de los países participantes en función de su desempeño general. Algunos autores denominan como “PISA shock” a la tendencia de los gobiernos a modificar sus políticas educativas como reacción a los resultados de PISA, habitualmente basándose en los resultados descriptivos iniciales y no en los análisis secundarios más profundos que se realizan de los datos (Wiseman, 2013). Sin embargo, estos resultados pueden llevar a la realización de inferencias equivocadas sobre la calidad de los sistemas educativos (Jornet, 2016)

Por otro lado, aunque la recogida de datos es cíclica, la naturaleza transversal de la prueba supone una imposibilidad de establecer relaciones causales y de realizar un seguimiento de los efectos de diversos factores en la evolución del rendimiento de los sujetos a lo largo del tiempo (Rutkowski et al., 2017). Sin embargo, esta perspectiva temporal es vital para poder realizar un análisis sobre la efectividad de las diversas políticas de un sistema educativo a lo largo del tiempo.

Debido a estas dos cuestiones, el presente artículo propone un análisis secundario de los datos de PISA con un enfoque de estudio de tendencia con un doble objetivo. Por un lado, ofrecer un análisis multivariante y multinivel más profundo que los resultados de PISA que habitualmente trascienden a la población general y a las personas encargadas de la toma de decisiones políticas, y por otro lado, obtener una perspectiva temporal que permita analizar la evolución del efecto de los factores contextuales en el rendimiento del alumnado español a través de dos ediciones distintas de PISA.

### **Factores asociados al rendimiento**

Aunque el objetivo principal de las evaluaciones a gran escala es estudiar el rendimiento de los y las estudiantes participantes en determinadas competencias o destrezas, el análisis secundario más habitualmente realizado es el estudio de los factores de diversa naturaleza que están relacionados con dicho rendimiento.

En los años 90 tuvo su auge la investigación en torno a la eficacia escolar. Este hecho trajo consigo el desarrollo de diferentes modelos para desarrollar estas investigaciones entre los que destaca el modelo dinámico planteado por Creemers y Kyriakides . Según este modelo existen diferentes niveles de factores que explican el rendimiento educativo del estudiantado, desde el nivel nacional hasta el nivel de alumno/a, pasando por la escuela, el profesorado (o la clase) o las características contextuales del propio estudiantado (Figura 1). Siguiendo el modelo indicado, el más empleado hasta la fecha, nuestra investigación se centra en los factores a nivel de alumno/a y su contexto y a nivel de escuela que demuestran estar interrelacionados y nos sirve de apoyo a la hora de definir la naturaleza de la eficacia escolar (Kyriakides et al., 2009).

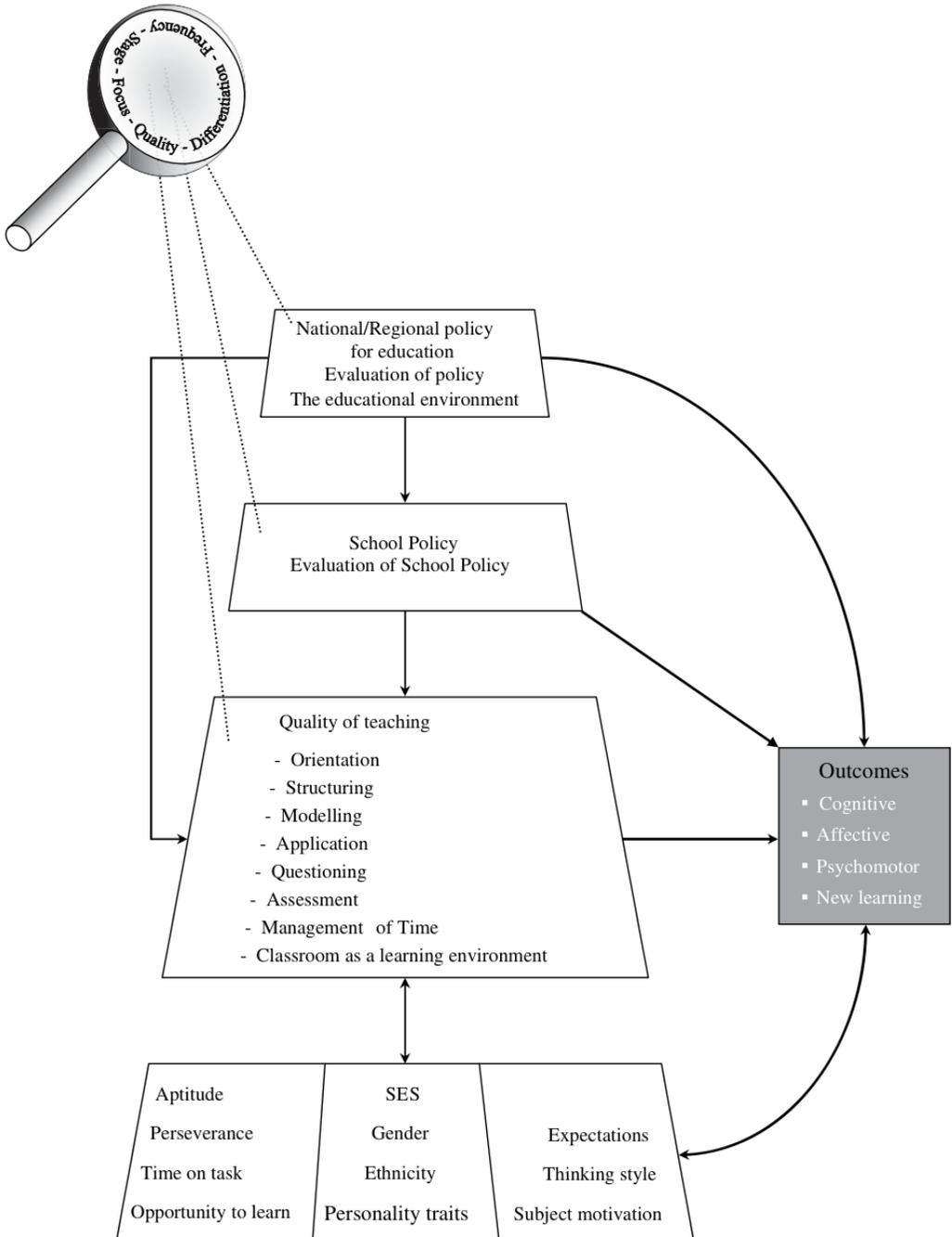


Figura 1. Modelo dinámico de la eficacia escolar (Kyriakides et al., 2009, p. 13)

El conjunto de factores que está significativamente relacionado con el rendimiento de los y las estudiantes varía ampliamente en función de la competencia sobre la que se realice el estudio o la metodología de investigación empleada, así como en función del país o entorno cultural del que provengan los y las estudiantes. Sin embargo, hay ciertas variables, fundamentalmente de carácter sociodemográfico y de contexto educativo, que consistentemente demuestran su destacable influencia sobre el rendimiento del alumnado en edad escolar.

Dada la naturaleza de las evaluaciones a gran escala, que recogen datos a dos niveles (estudiante y centro educativo), podemos encontrar dos tipos principales de variables estudiadas basándonos en la revisión previa de la literatura.

Las variables a nivel de estudiante son las que suelen mostrar más influencia en su rendimiento, siendo las más frecuentes el nivel socioeconómico (Lenkeit, 2012; Pomianowicz, 2021; Yetişir y Bati, 2021), el género (Laukaityte y Rolfsman, 2020; Martínez-Abad et al., 2020; H. Wu et al., 2020), el estatus migratorio (Doncel Abad y Cabrera Álvarez, 2020; Gómez-Fernández y Mediavilla, 2021; Pomianowicz, 2021), haber repetido curso (Doncel Abad y Cabrera Álvarez, 2020; Autora et al., 2018; Martínez-Abad, 2019) o hablar en casa una lengua distinta a la lengua vehicular del sistema educativo (Doncel Abad y Cabrera Álvarez, 2020; Martínez-Abad et al., 2020; Pomianowicz, 2021).

Con respecto a los factores del centro educativo, la única variable que consistentemente muestra una influencia en el rendimiento de los y las estudiantes es el efecto contextual del nivel socioeconómico del estudiantado del centro (Ding y Homer, 2020; Kameshwara et al., 2020; H. Wu et al., 2020), aunque algunos estudios también encuentran relevantes otro tipo de variables como el tamaño del centro o su titularidad (Hu et al., 2018; Martínez-Abad, 2019).

## **Longitudinalidad y estudios de tendencia**

Una de las críticas que más frecuentemente se realizan con respecto a la configuración de PISA es su carácter transversal, que no permite analizar la evolución de los y las estudiantes y sus circunstancias ni establecer relaciones de carácter causal entre las variables de estudio (Carabaña, 2015; Autora, 2020; Han, 2018), o en su defecto de datos sobre el rendimiento previo de los y las estudiantes (Dumay y Dupriez, 2014; Willms, 2010), lo cual supone la imposibilidad de establecer modelos de valor añadido que muestren una información más ajustada de los efectos del centro educativo.

Las características inherentes a esta prueba de evaluación a gran escala no nos permiten realizar estudios longitudinales, caracterizados por hacer referencia a diferentes evaluaciones sobre los mismos sujetos a lo largo del tiempo (de Miguel, 1985). Así, teniendo en cuenta que el objetivo de este estudio es realizar conclusiones sobre la evolución de los resultados y características de una población concreta a lo largo del tiempo, existirían otras dos opciones metodológicas: un estudio por cohortes y un estudio de tendencia. En el caso del estudio de cohortes, aunque los sujetos de los que se recoge información no necesariamente son los mismos en cada medición, la población de la que se extraen estos sujetos sí es estable (p. ej. sujetos nacidos en un año concreto o personas de una misma promoción universitaria) (Bisquerra, 2004; Cohen et al., 2017). Sin embargo, teniendo en

cuenta que la población objetivo de PISA son el estudiantado que tienen 15 años en el momento de la evaluación, esta metodología tampoco es viable.

Como alternativa ante la imposibilidad de desarrollar un estudio longitudinal o de cohortes, esta investigación propone la realización de un estudio de tendencias. Los estudios de tendencias suponen una variante del estudio de cohortes, en el que se pretende describir un cambio mediante la medición de poblaciones que no se mantienen estables a lo largo del tiempo, y por tanto los sujetos que componen la población objeto de estudio son diferentes en cada recogida de datos (Bisquerra, 2004; Cohen, 2017; Sullivan y Calderwood, 2017).

Aunque al variar de muestra entre las distintas ediciones de PISA se introducen elementos que pueden distorsionar los resultados, los estudios de tendencias suponen una estrategia razonablemente correcta capaz de identificar las corrientes de un comportamiento mediante el análisis de informaciones recogidas en distintos momentos temporales (Bisquerra, 2004).

Teniendo en cuenta todo esto, el objetivo general planteado para este estudio es la comparación del impacto de las principales variables sociodemográficas en el rendimiento competencial de las y los estudiantes españoles entre las evaluaciones PISA 2015 y 2018 para valorar su evolución.

## **Método**

Este estudio presenta un análisis de datos secundario de dos ciclos distintos de PISA basado en un diseño *ex post facto*, ya que se realiza sobre datos previamente recogidos y no implica ningún tipo de manipulación de variables. Aunque cada una de las recogidas de datos tiene un carácter transversal, al aunar el análisis de ambas ediciones se propone un estudio de tendencias, como se justificó en el anterior apartado.

### **Población y Muestra**

La población del estudio está compuesta por las y los estudiantes españoles que tenían 15 años en 2015 y 2018. La selección de estos ciclos de PISA viene motivada por las variables disponibles en los cuestionarios de contexto, las cuales sirven como variables predictoras en el presente estudio. Las ediciones de 2015 y 2018 contienen las mismas variables, sin embargo, en las ediciones anteriores a 2015 faltan algunas variables que han demostrado ser relevantes a la hora de predecir el rendimiento, como el número de cambios de centro que un/a estudiante ha realizado a lo largo de su historial escolar o la duración de la escolarización en el periodo anterior a la educación obligatoria (Educación Infantil) (Autora et al., 2018). Por este motivo, y para garantizar el grado más alto de comparabilidad posible, se decidió seleccionar sólo los datos procedentes de las evaluaciones de 2015 y 2018.

La muestra se obtiene a través de un proceso de muestreo bietápico, en el que inicialmente se seleccionan los centros educativos que van a participar (representativos en función de la región/país y la titularidad de los centros) y después se determina qué estudiantes deberán realizar las pruebas dentro de cada centro educativo (OECD, 2019).

En el caso de España, la muestra recogida es representativa a nivel de Comunidad Autónoma, lo que permite una desagregación de los datos por regiones que no está disponible en todos los países participantes y que facilita la comparación inter-regional.

Ya que uno de los objetivos del estudio es analizar el efecto de factores a nivel de centro educativo, se han excluido los centros con menos de 20 alumnos encuestados, siguiendo la recomendación de algunos autores (Joaristi et al., 2014; Martínez-Abad, 2019; Martínez-Abad et al., 2017; Meunier, 2011). Así, la muestra del estudio queda conformada por un total de 65684 estudiantes y 1873 centros educativos, distribuidos por año como se indica en la tabla 1.

Tabla 1

*Muestra del estudio, por ciclo de PISA y nivel*

	2015	2018
Alumnado	31273	34411
Centros educativos	897	976

## **Instrumentos**

El presente estudio utiliza información recogida a través de dos tipos de instrumentos, siendo ambos parte de la evaluación PISA<sup>2</sup>. En primer lugar, se hace uso de los resultados de los cuestionarios de evaluación de competencias (matemática, científica y lectora). Estos cuestionarios están conformados por diversos tipos de ítems, en función del contenido concreto al que se refieren, el nivel de dificultad o el tipo de respuesta que requieren (múltiple, cerrada, abierta).

Por su parte, los cuestionarios de contexto, aplicados a alumnado y centros, están dirigidos a recabar información sobre resultados no cognitivos (autoeficacia, motivación, actitudes hacia el centro y el aprendizaje), condiciones individuales (entorno educativo y socioeconómico), y las características procesuales y organizativas del entorno institucional (estructura, recursos y procesos del centro educativo). Estos cuestionarios tienen un carácter auto informado, y son respondidos por el alumnado y el equipo directivo de los centros participantes.

## **Variables**

Las variables criterio utilizadas en este estudio son las puntuaciones de los estudiantes participantes en cada una de las competencias principales evaluadas: competencia matemática, competencia científica y competencia en comprensión lectora<sup>3</sup>.

2 Las bases de datos de PISA son de consulta pública y pueden encontrarse en el siguiente enlace: <https://www.oecd.org/pisa/data/>

3 Para una definición en profundidad de estas variables, puede consultarse el Informe Español de PISA 2018 (Ministerio de Educación y Formación Profesional, 2019).

Por otro lado, las variables predictoras incluidas en el estudio son algunas de las variables de carácter contextual a nivel de estudiante y centro educativo reflejadas en los cuestionarios de contexto de PISA (tabla 2) y basadas en la revisión de la literatura expuesta con anterioridad.

Tabla 2

*Variables contextuales de los modelos multinivel*

	Variable	Etiqueta		Rango
		PISA 2015	PISA 2018	
Nivel 1 - Alumnos	Género	ST004D01T	ST004D01T	0: Masculino 1: Femenino
	Mes de nacimiento (2015)	ST003D02T		1-12
	Edad (2018)		AGE	Continua
	Curso	GRADE	GRADE	1º ESO – 1º Bachillerato
	Índice socio-económico y cultural (ESCS)	ESCS	ESCS	Continua
	Estatus migratorio	IMMIG	IMMIG	0: Nativo, 1: Inmigrante de 2ª generación 2: Inmigrante de 1ª generación
	Condición de repetidor	REPEAT	REPEAT	0: No 1: Sí
	Número de cambios de centro educativo (2015)	SCCHANGE		0: Ningún cambio 1: Un cambio 2: Dos o más cambios
	Número de cambios de centro educativo (2018)		SCCHANGE	Continua
	Número de años cursados en Educación Infantil	DURECEC	DURECEC	Continua
	Idioma hablado en el hogar	ST022Q01TA	ST022Q01TA	0: Idioma de la prueba 1: Otro idioma

Nivel 2 - Centros	Tamaño del centro	SCHSIZE	SCHSIZE	Continua
	Tamaño de las clases	CLSIZE	CLSIZE	Continua
	Escasez de recursos	EDUSHORT	EDUSHORT	Continua
	Escasez de profesorado	STAFFSHORT	STAFFSHORT	Continua
	Titularidad del centro	SCHLTYPE	SCHTYPE	1: Privado 2: Concertado 3: Público
	Ratio profesor-alumnos	STRATIO	STRATIO	Continua
	Localización del centro	SC001Q01TA	SC001Q01TA	1: Área rural (menos de 3000 habitantes) 2: Pueblo pequeño (3 000-15 000 habitantes) 3: Pueblo (15 000-100 000 habitantes) 4: Ciudad (100 000 a 1 000 000 habitantes) 5: Ciudad grande (Más de 1 000 000 habitantes)
	Proporción de docentes con Máster	PROAT5AM	PROAT5AM	Continua
	Proporción de docentes con Doctorado	PROAT6	PROAT6	Continua
	ESCS medio*	ST_ESCS	ST_ESCS	Continua
	Proporción de repetidores*	ST_REPEAT	ST_REPEAT	Continua
	Proporción de alumnado inmigrantes*	ST_INM	ST_INM	Continua
	Proporción de alumnas*	ST_GEN	ST_GEN	Continua

*Nota: \*Variable agregada a nivel de centro desde la base de datos de estudiantes seleccionadas por su relevancia demostrada en investigaciones previas.*

Las variables categóricas con más de dos opciones de respuesta (estatus migratorio, cambios de centro educativo 2015, titularidad del centro...) fueron convertidas en variables *dummy* para favorecer la claridad del análisis.

## Análisis de datos

La técnica de análisis de datos utilizada es la regresión multinivel, también conocida como Modelos Jerárquicos Lineales. La elección de este tipo de análisis está motivada fundamentalmente por la naturaleza de los datos procedentes de las evaluaciones a gran escala, los cuales presentan una estructura anidada en dos niveles (estudiantes dentro de escuelas), suponiendo por tanto la existencia de variabilidad a ambos niveles. El uso de la regresión multinivel permite considerar simultáneamente el efecto de variables de diferentes niveles, evitando así los sesgos que podrían provenir de utilizar técnicas de agregación o desagregación de los datos para poder llevar a cabo análisis de regresión simple (Gaviria Soto y Castro Morera, 2005; Raudenbush y Bryk, 2002; Snijders y Bosker, 2011).

El modelo planteado responde a un diseño de efectos mixtos, con intercepto aleatorio y pendientes fijas. El intercepto aleatorio permite que las unidades de nivel más alto en la jerarquía del modelo (los centros educativos en este caso) tengan diferente media en la variable dependiente o criterio (Hayes, 2006). Por su parte, los efectos de las covariables en este caso se mantienen fijos entre los distintos centros educativos. Un modelo de pendientes aleatorias, en el que se permite que las covariables tengan una pendiente diferente en cada uno de los centros, supondría un reflejo más fiel de la realidad educativa. Sin embargo, introduce también un nivel de complejidad al análisis, y dado que es especialmente recomendable cuando el objetivo del estudio es realizar análisis de eficacia diferencial de los centros (Clarke et al., 2010), no se estima necesario en el presente caso. Teniendo en cuenta que el coeficiente de regresión obtenido para los modelos de ambas ediciones no se ofrece estandarizado, los resultados se interpretarán a partir del estadístico *t*, que sí se encuentra estandarizado en función del error típico (De Veaux et al., 2021).

Para poder comprobar el cumplimiento de los supuestos previos de la regresión multinivel y evaluar la idoneidad de este tipo de análisis, se calcula el Coeficiente de Correlación Intraclase (ICC) del modelo nulo (sin covariables) de las variables criterio, el cual indica la proporción de varianza del rendimiento de los y las estudiantes que es atribuible al segundo nivel de análisis (centros) (Autora et al., 2018). El valor de este coeficiente debe ser igual o superior al 10% para considerar adecuada la aplicación de esta metodología (Lee, 2000), cuestión que se cumple para todas las variables criterio utilizadas.

Tabla 3

*Coefficiente de Correlación Intraclase del modelo nulo, por ciclo y competencia*

	2015			2018		
ICC	Matemáticas	Ciencias	Lectura	Matemáticas	Ciencias	Lectura
	12.26%	12.41%	12.04%	13.14%	11.39%	14.21%

El análisis fue llevado a cabo con el programa estadístico especializado en regresión multinivel HLM. Este programa permite el tratamiento de los datos procedentes de evaluaciones a gran escala con el cumplimiento de todas las garantías estadísticas indi-

casas para este tipo de datos (OECD, 2019). En primer lugar, se permite el tratamiento de variables criterio con valores plausibles, como es el caso de los datos de rendimiento de PISA. Los valores plausibles pueden definirse como el rango de habilidades que un/a estudiante podría tener razonablemente, dadas sus respuestas a los ítems (Wu, 2005), y son particularmente útiles en el análisis de datos con diseños complejos como PISA porque facilitan el abordaje de los problemas relacionados con sesgos en la estimación de los parámetros poblacionales y facilitan también la computación de los errores típicos de estimación en diseños muestrales complejos.

Por otro lado, el programa HLM permite también el uso simultáneo de pesos muestrales a nivel de estudiante y de centro educativo. Estas ponderaciones reflejan la probabilidad desigual de los y las estudiantes y centros de ser seleccionados en el proceso de muestreo, y su aplicación garantiza la correcta representatividad de la muestra.

## Resultados

A continuación, se presentan los resultados del proceso de modelización multinivel para cada una de las competencias.

### Matemáticas

Una vez seleccionadas solamente aquellas variables con una relación significativa con el rendimiento en la competencia matemática, los modelos resultantes responden a las siguientes ecuaciones (1 y 2).

*Matemáticas 2015*<sub>ij</sub>

$$\begin{aligned} &= \gamma_{00} + \gamma_{01} * ESCS \text{ nivel } 2_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} \\ &* \text{Inmigrante } 1^{\text{a}} \text{ generación}_{ij} + \gamma_{30} * \text{Curso}_{ij} + \gamma_{40} \\ &* \text{Repetición}_{ij} + \gamma_{50} * ESCS_{ij} + \gamma_{60} * \text{Cambio centro } 1_{ij} + \gamma_{70} \\ &* \text{Cambio centro } 2_{ij} + u_{0j} + r_{ij} \end{aligned}$$

Ecuación 1

*Matemáticas 2018*<sub>ij</sub>

$$\begin{aligned} &= \gamma_{00} + \gamma_{01} * \text{Escasez personal}_j + \gamma_{02} * ESCS \text{ nivel } 2_j + \gamma_{10} \\ &* ESCS_{ij} + \gamma_{20} * \text{Género}_{ij} + \gamma_{30} * \text{Curso}_{ij} + \gamma_{40} * \text{Edad}_{ij} + \gamma_{50} \\ &* \text{Lengua}_{ij} + \gamma_{60} * \text{Repetición}_{ij} + \gamma_{70} * \text{Cambio centro}_{ij} + u_{0j} \\ &+ r_{ij} \end{aligned}$$

Ecuación 2

Donde:

*Matemáticas*<sub>ij</sub> es la puntuación media del centro en matemáticas;

$\gamma_{00}$  es la media de todos los centros en matemáticas;

$\gamma_{01} - \gamma_{02}$  son las covariables de nivel 2;

$\gamma_{10} - \gamma_{70}$  son las covariables de nivel 1;

$u_{0j}$  es la diferencia entre la puntuación del centro en matemáticas y la media general de todos los centros; y

$r_{ij}$  es el residuo del nivel 1.

Los coeficientes correspondientes a cada una de las covariables, junto con su valor de t-ratio y su nivel de significación pueden consultarse en la tabla 4.

Tabla 4

*Resultados de la regresión multinivel para la competencia matemática, 2015 y 2018*

Efecto fijo	2018			2015		
	Coef.	t-ratio	Sig.	Coef.	t-ratio	Sig.
INTERCEPTO	375.84	8.831	<0.001	543.069	313.438	<0.001
N2 Escasez personal	-2.635	-2.264	0.024			
N2 ESCS medio	14.864	6.860	<0.001	13.000	6.820	<0.001
N1 ESCS	11.624	12.776	<0.001	9.787	8.729	<0.001
N1 Género	-18.223	-9.225	<0.001	-22.578	-11.474	<0.001
N1 Curso	26.800	5.844	<0.001	34.859	9.048	<0.001
N1 Edad	10.050	3.709	<0.001			
N1 Lengua	-6.396	-2.863	0.006			
N1 Repetición	-55.031	-8.158	<0.001	-35.803	-7.243	<0.001
N1 Cambios de centro	-9.499	-9.643	<0.001			
N1 Cambio de centro (1)				-10.649	-5.724	<0.001
N1 Cambio de centro (2+)				-15.602	-4.724	<0.001
N1 Inmigrante 1ª gen				-11.806	-2.433	0.026

La mayoría de las variables significativas para el rendimiento en matemáticas de las y los estudiantes españoles se mantienen entre ambas ediciones de la evaluación. Sin embargo, hay algunos cambios, tanto a nivel de las variables que se incluyen como en la magnitud de la influencia de estas en el rendimiento.

En primer lugar, cabe destacar que las variables relacionadas con el nivel socioeconómico, tanto a nivel individual como a nivel de centro, mantienen coeficientes parecidos en ambas evaluaciones, por lo que parece que su influencia en el rendimiento es estable.

En el apartado de variables cuyo efecto ha reducido su magnitud entre 2015 y 2018 se encuentran el género, cuya influencia negativa para las mujeres ha disminuido ligeramente, y el curso, cuyo impacto para aquellos estudiantes que se encuentran en cursos más bajos que lo que les correspondería por edad ha disminuido considerablemente. La variable relativa al efecto generado por la condición de ser inmigrante de primera generación merece una mención aparte, ya que esta ha desaparecido como variable relevante para el modelo entre ambas ediciones.

Por otro lado, la variable relacionada con la repetición de curso ha aumentado su efecto negativo sobre el alumnado repetidor.

Por último, en la edición de 2018 aparecen algunas variables nuevas que no habían sido significativas en 2015 (escasez de personal, edad del estudiante y lengua hablada en el hogar), aunque ninguna de ellas presenta un efecto muy relevante ni en términos de coeficiente ni de importancia relativa para el modelo (t-ratio).

## Ciencias

Tras el proceso de análisis de datos y la selección de variables significativas, los modelos explicativos de la competencia científica del alumnado español quedan configurados según las siguientes ecuaciones (3 y 4).

$$\begin{aligned}
 \text{Ciencia 2015}_{ij} = & \gamma_{00} + \gamma_{01} * \text{Tamaño centro}_j + \gamma_{02} * \text{Escasez personal}_j + \gamma_{03} \\
 & * \text{ESCS nivel 2}_j + \gamma_{04} * \text{Porcentaje de repetidores}_j + \gamma_{05} \\
 & * \text{Porcentaje de alumnas}_j + \gamma_{10} * \text{Mes nacimiento}_{ij} + \gamma_{20} \\
 & * \text{Género}_{ij} + \gamma_{30} * \text{Inmigrante 1}^{\text{a}} \text{ generación}_{ij} + \gamma_{40} * \text{Curso}_{ij} \\
 & + \gamma_{50} * \text{Repetición}_{ij} + \gamma_{60} * \text{ESCS}_{ij} + \gamma_{70} * \text{Cambio centro 1}_{ij} \\
 & + \gamma_{80} * \text{Cambio centro 2}_{ij} + u_{0j} + r_{ij}
 \end{aligned}$$

Ecuación 3

$$\begin{aligned}
 \text{Ciencia 2018}_{ij} = & \gamma_{00} + \gamma_{01} * \text{ESCS}_j + \gamma_{10} * \text{ESCS}_{ij} + \gamma_{20} * \text{Género}_{ij} + \gamma_{30} * \text{Curso}_{ij} \\
 & + \gamma_{40} * \text{Edad}_{ij} + \gamma_{50} * \text{Lengua}_{ij} + \gamma_{60} * \text{Repetición}_{ij} + \gamma_{70} \\
 & * \text{Cambio centro}_{ij} + u_{0j} + r_{ij}
 \end{aligned}$$

Ecuación 4

Donde:

$\text{Ciencia}_{ij}$  es la puntuación media del centro en ciencia;

$\gamma_{00}$  es la media de todos los centros en ciencia;

$\gamma_{01}$ - $\gamma_{05}$  son las covariables de nivel 2;

$\gamma_{10}$  -  $\gamma_{70}$  son las covariables de nivel 1;

$u_{0j}$  es la diferencia entre la puntuación del centro en ciencia y la media general de todos los centros; y

$r_{ij}$  es el residuo del nivel 1.

Los coeficientes y niveles de significación de estas variables pueden consultarse en la tabla 5.

Tabla 5

Resultados de la regresión multinivel para la competencia científica, 2015 y 2018

Efecto Fijo	2018			2015		
	Coef.	t-ratio	Sig.	Coef.	t-ratio	Sig.
INTERCEPTO	440.24	11.360	<0.001	537.68	73.71	<0.001
N2 ESCS medio	12.378	4.960	<0.001	20.75	7.96	<0.001
N2 Tamaño centro				-0.01	-2.24	0.025

N2 Escasez personal				2.49	2.14	0.032
N2 Porcentaje repetidores				31.95	3.35	<0.001
N2 Porcentaje alumnas				31.84	2.66	0.008
N1 ESCS	10.630	10.978	<0.001	9.78	13.99	<0.001
N1 Género	-13.174	-7.083	<0.001	-19.76	-11.70	<0.001
N1 Curso	23.026	6.060	<0.001	38.29	13.25	<0.001
N1 Edad/Mes nacimiento	6.040	2.537	0.013	-0.59	-2.56	0.011
N1 Lengua	-10.410	-3.751	<0.001			
N1 Repetición	-53.281	-10.756	<0.001	-34.29	-8.86	<0.001
N1 Cambios de centro	-8.751	-7.377	<0.001			
N1 Cambio de centro (1)				-10.36	-5.34	<0.001
N1 Cambio de centro (2+)				-16.96	-6.43	<0.001
N1 Inmigrante 1ª gen				-9.63	-3.34	<0.001

Al igual que en la competencia matemática, se observan algunas variables que se mantienen entre ambas ediciones, y otras muchas que varían.

De nuevo las variables centradas en el nivel socioeconómico tanto a nivel de estudiante como de centro se mantienen. Aunque el nivel socioeconómico de los y las estudiantes presenta un impacto parecido en 2015 y 2018, parece que el impacto del efecto contextual del ESCS (a nivel de centro) ha disminuido considerablemente entre ambas ediciones.

Del resto de variables que se repiten, algunas han reducido también su impacto, como, por ejemplo, el género. Esta variable muestra una reducción en el impacto negativo para las alumnas, disminuyéndose de esta manera la brecha de género en la competencia científica. El impacto del curso también ha disminuido, acercando más en puntuaciones a aquellos/as estudiantes que se encuentran en cursos diferentes en el momento de la evaluación.

Por otro lado, vuelve a resultar destacable el aumento del impacto de la repetición de curso en el rendimiento del estudiantado.

Como puede observarse, hay numerosas variables que desaparecen entre ambas ediciones, como el tamaño del centro, la escasez de personal, el porcentaje de repetidores o el porcentaje de alumnas a nivel de centro, o la condición de ser inmigrante de primera generación a nivel de estudiante.

En el análisis de la edición de 2018 aparecen algunas variables nuevas, como la lengua hablada en el hogar, con un impacto relativamente bajo y negativo en el ren-

dimiento de aquellos/as estudiantes que no hablan la lengua vehicular del sistema educativo en su hogar.

### Lectura

Tras el proceso de modelización, y una vez que se han eliminado todas las variables que no presentan una relación significativa con el rendimiento del estudiante en comprensión lectora, los modelos quedan configurados según las siguientes ecuaciones (5 y 6).

$$\begin{aligned}
 \text{Lectura 2015}_{ij} = & \gamma_{00} + \gamma_{01} * \text{ESCS medio}_j + \gamma_{02} * \text{Porcentaje de repetidores}_j \\
 & + \gamma_{03} * \text{Porcentaje de alumnas}_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} * \text{Curso}_{ij} \\
 & + \gamma_{30} * \text{Repetición}_{ij} + \gamma_{40} * \text{ESCS}_{ij} + \gamma_{50} * \text{Lengua}_{ij} + \gamma_{60} \\
 & * \text{Cambio centro 1}_{ij} + \gamma_{70} * \text{Cambio centro 2}_{ij} + u_{0j} + r_{ij}
 \end{aligned}$$

Ecuación 5

$$\begin{aligned}
 \text{Lectura 2018}_{ij} = & \gamma_{00} + \gamma_{01} * \text{ESCS medio}_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} * \text{Curso}_{ij} + \gamma_{30} \\
 & * \text{Repetición}_{ij} + \gamma_{40} * \text{ESCS}_{ij} + \gamma_{50} * \text{Lengua}_{ij} + \gamma_{60} \\
 & * \text{Inmigrante de 1ª generación}_{ij} + \gamma_{70} * \text{Cambio centro}_{ij} + \gamma_{80} \\
 & * \text{Edad}_{ij} + u_{0j} + r_{ij}
 \end{aligned}$$

Ecuación 6

Donde:

$Lectura_{ij}$  es la puntuación media del centro en lectura;

$\gamma_{00}$  es la media de todos los centros en lectura;

$\gamma_{01}$ - $\gamma_{03}$  son las covariables de nivel 2;

$\gamma_{10}$  - $\gamma_{80}$  son las covariables de nivel 1;

$u_{0j}$  es la diferencia entre la puntuación del centro en lectura y la media general de todos los centros; y

$r_{ij}$  es el residuo del nivel 1.

Los coeficientes y niveles de significación de estas variables pueden consultarse en la tabla 6.

Tabla 6

Resultados de la regresión multinivel para la competencia en comprensión lectora, 2015 y 2018

Efecto Fijo	2018			2015		
	Coef.	t-ratio	Sig.	Coef.	t-ratio	Sig.
INTERCEPTO	396.46	11.310	<0.001	522.475	84.912	<0.001
N2 ESCS medio	14.353	6.114	<0.001	20.293	7.633	<0.001

N2 Porcentaje repetidores				36.502	3.371	<0.001
N2 Porcentaje alumnas				24.719	2.338	0.020
N1 ESCS	10.167	12.998	<0.001	8.299	8.748	<0.001
N1 Género	15.442	10.362	<0.001	6.776	3.665	<0.001
N1 Curso	23.543	7.360	<0.001	41.092	11.757	<0.001
N1 Edad	7.542	3.416	<0.001			
N1 Lengua	-10.475	-4.387	<0.001	-6.718	-2.220	0.033
N1 Repetición	-52.633	-12.657	<0.001	-30.72	-6.09	<0.001
N1 Cambios de centro	-8.953	-12.230	<0.001			
N1 Cambios centro (1)				-8.857	-4.164	<0.001
N1 Cambios centro (2+)				-20.642	-6.403	<0.001

Con respecto a las variables socioeconómicas, se repite el resultado de los dos modelos anteriores: mientras que el efecto de la variable a nivel individual se mantiene (y en este caso incluso aumenta), su efecto contextual a nivel de centro disminuye.

Otra variable cuyo efecto disminuye entre ambas ediciones es el curso.

Por otro lado, se vuelve a repetir el fenómeno relacionado con la variable repetición, cuyo efecto aumenta a casi el doble en función de la t-ratio. En el caso de la comprensión lectora, al contrario que en los dos anteriores, el efecto del género aumenta considerablemente, ensanchando así la brecha de género, que en esta competencia funciona a favor de las alumnas.

En el caso de la competencia en comprensión lectora, sólo dos variables desaparecen entre la edición de 2015 y la de 2018: porcentaje de repetidores/as y porcentaje de alumnas. Por otro lado, aparece como nueva variable la edad del alumno/a.

### Estudio descriptivo de la brecha de género

Debido a los resultados obtenidos en los tres modelos, los cuales evidencian cambios en la brecha de género entre alumnos y alumnas (una reducción en el caso de las competencias matemática y científica y un aumento en el caso de la comprensión lectora), cabe realizar un análisis descriptivo para comprobar si la brecha se ha estrechado (o aumentado, en el caso de lectura) porque las alumnas han mejorado o porque los alumnos han empeorado. Para contestar a esta cuestión, se plantea un análisis de medias en función del género del estudiantado entre ambas competencias para comprobar cuál ha sido la evolución de los niveles competenciales. Los resultados de este análisis se pueden ver en la tabla 7.

Tabla 7

*Medias de las puntuaciones en las tres competencias en función del género*

Competencia	Alumnas			Alumnos		
	2018	2015	Dif. 2018-2015	2018	2015	Dif. 2018-2015
Mate	487,96	487,47	0,48	495,76	499,22	-3,46
Ciencias	489,63	495,84	-6,22	493,12	503,57	-10,45
Lectura	495,76	511,06	-15,29	471,25	493,06	-21,81

Como se puede observar, el rendimiento de los y las estudiantes en general ha empeorado entre 2015 y 2018 en todas las competencias. Por tanto, la clave es la magnitud del descenso, que en todos los casos ha sido menor para las alumnas, y especialmente en el ámbito de la comprensión lectora.

### **Coefficientes de correlación intraclass de los modelos finales**

Para completar el análisis multinivel cabe volver a explorar el índice de correlación intraclass, en esta ocasión incluyendo las covariables significativas en cada modelo, para poder comprobar la proporción de variabilidad al nivel del centro que ha conseguido explicar cada modelo (tabla 8).

Tabla 8

*Coefficiente de Correlación Intraclass del modelo nulo y el modelo final, por ciclo y competencia.*

	2015			2018		
	Mate	Ciencias	Lectura	Mate	Ciencias	Lectura
ICC nulo	12.26%	12.41%	12.04%	13.14%	11.39%	14.21%
ICC final	4.55%	5.6%	5.07%	6.15%	6.86%	10.47%
Proporción explicada	62.89%	54.88%	57.89%	52.05%	39.77%	26.32%

Estos datos ilustran la gran relevancia que tienen las variables sociodemográficas y de contexto educativo a la hora de explicar la variabilidad de las variables criterio presentada a nivel de centro educativo, llegando a explicar en algunos modelos hasta casi un 60%.

### **Discusión**

Aunque cada modelo presenta sus particularidades, se pueden encontrar algunas características que se observan consistentemente en los tres.

En primer lugar, el efecto contextual del nivel socioeconómico, es decir, la media de esta variable en cada uno de los centros educativos estudiados, ha disminuido consider-

ablemente, lo cual es un aspecto positivo de evolución del sistema educativo, de acuerdo con lo observado en otras investigaciones (Lenkeit, 2012; Sirin, 2005).

En segundo lugar, una cuestión llamativa es el aumento del efecto de la repetición en el rendimiento de las tres competencias estudiadas, como en el estudio planteado por Martínez-Abad (2019). Sin embargo, otra de las características comunes a los tres modelos podría ayudar a explicar esta cuestión, al menos en parte. Al igual que el efecto de la repetición ha aumentado consistente y considerablemente, la variable curso ha visto reducido su impacto a un nivel notable en las tres competencias. Teniendo en cuenta que el hecho de que un sujeto haya repetido o no y en qué curso se encuentra en el momento de la evaluación son variables altamente relacionadas, podría hipotetizarse que, por algún motivo, parte del efecto asignado al curso en 2015 se ha trasladado a la repetición de curso en 2018, aunque serían necesarios más estudios para profundizar en esta cuestión.

Otra cuestión que cabe analizar es la brecha de género en el rendimiento de los y las participantes. En las dos competencias de ámbito científico (ciencias y matemáticas) esta brecha, que favorece a los estudiantes varones, se ha reducido notablemente (entre un 30% y un 50% tanto en términos de relevancia relativa para el modelo), cuestión que se encuentra alineada con los resultados de otros estudios recientes (Molina Portillo et al., 2022), que también indican una reducción de la brecha de género en materias STEM. Sin embargo, no es este el caso de la comprensión lectora, en la cual la brecha de género, que en este caso favorece a las alumnas, ha aumentado su efecto en más del doble entre las ediciones de 2015 y 2018.

Otra variable destacable es la condición de inmigrante de primera generación. En 2015, esta variable formaba parte de los modelos de ciencias y matemáticas. En 2018, este impacto ha desaparecido de estos dos modelos, igualando a los y las estudiantes inmigrantes de primera generación con el resto de sus compañeros/as en cuanto a rendimiento en estas competencias. Sin embargo, mientras que en el modelo de comprensión lectora en 2015 no se encontraba esta variable entre las significativas, en 2018 sí ha aparecido en este modelo y con una importancia relativa media-alta.

## **Conclusiones**

Aunque existen numerosos estudios dedicados a analizar el impacto de las variables sociodemográficas y educativas en el rendimiento del estudiantado a través de las evaluaciones a gran escala, la naturaleza transversal de dichas evaluaciones hace que sea complicado valorar la evolución de los sistemas educativos en cuanto a su capacidad para controlar los efectos del contexto socioeconómico. Por este motivo, en este estudio se plantea una comparativa entre las dos últimas ediciones de PISA con el objetivo de ganar una cierta perspectiva temporal que ayude a valorar la eficacia del sistema educativo a lo largo del tiempo.

En general, las conclusiones con respecto a la evolución del sistema educativo en su capacidad para reducir las desigualdades educativas son positivas. A través del estudio comparativo se ha detectado la reducción del impacto de variables relevantes como el nivel socioeconómico medio de los y las estudiantes de un centro o la condición de ser inmigrante de primera generación, así como la reducción de la brecha de género en las competencias relacionadas con el ámbito STEM. Sobre este punto, teniendo en cuenta

la información contenida en la tabla 7, se puede concluir que la reducción de la brecha de género en las materias STEM y el ensanchamiento de la brecha en la comprensión lectora se debe a un menor descenso en las puntuaciones de las alumnas entre ambas ediciones con respecto al descenso producido en el caso de los alumnos.

Cabe comentar alguna de las limitaciones de este estudio, que deben ser tenidas en cuenta a la hora de valorar los resultados y las conclusiones. En primer lugar, al tomar como base del análisis los datos procedentes de la evaluación PISA hay que asumir necesariamente las limitaciones propias de esta fuente, como son la falta de datos a nivel de aula (Lafontaine et al., 2015; Scheerens et al., 2015), lo cual impide estudiar variables relevantes como la metodología docente o el agrupamiento de los y las estudiantes, la naturaleza transversal de los datos (Autora, 2020; Yetişir y Bati, 2021), o las críticas existentes a la configuración de los cuestionarios de contexto (Autora et al., 2018; Li, 2016), lo cual afectaría fundamentalmente a la fiabilidad de las variables de carácter sociodemográfico y educativo. Por otro lado, el diseño del estudio también presenta algunas limitaciones, como la reducida muestra de ediciones estudiadas (solamente dos) debido a la falta de comparabilidad con las ediciones anteriores a 2015.

Los resultados y las conclusiones realizadas indican algunas líneas de investigación que sería interesante tomar en el futuro de cara a esclarecer algunas cuestiones que escapan del alcance de este estudio. Sería conveniente hacer un estudio estadístico más profundo del comportamiento y la evolución de las variables curso y repetición para comprobar la transferencia de efectos sugerida por este estudio. Por otro lado, el comportamiento reverso de la brecha de género en las materias STEM y en comprensión lectora merece ser estudiada con mayor detenimiento, para examinar posibles variables que puedan explicar dicho comportamiento y planificar políticas y prácticas educativas dirigidas a potenciar el rendimiento del estudiantado en todas las competencias independientemente de su género. Por último, otra cuestión sobre la que sería interesante indagar de cara a establecer políticas y prácticas educativas que fomenten la igualdad educativa es qué motivos hay detrás del descenso en el efecto ejercido por variables como el nivel socioeconómico medio de los estudiantes o la condición de ser inmigrante de primera generación.

### **Financiación**

Esta investigación forma parte del Proyecto PID2021-125775NB-I00 financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa. Del mismo modo, la investigación se realiza gracias a una Ayuda para contratos predoctorales para la formación de doctores con referencia PRE2019-087412 financiada por el Ministerio de Ciencia e Innovación del Gobierno de España.

### **Referencias**

- Bisquerra, R. (2004). *Metodología de la investigación educativa*. La Muralla.  
Carabaña, J. (2015). *La inutilidad de PISA para las escuelas*. Los Libros de la Catarata, D.L.

- Creemers, B. P. M. y Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. <https://doi.org/10.4324/9780203939185>
- Cohen, L., Manion, L., y Morrison, K. (2017). Surveys, longitudinal, cross-sectional and trend studies. In *Research methods in education* (pp. 334-360). Routledge. <https://doi.org/10.4324/9781315456539>
- de Miguel, M. (1985). Estrategias metodológicas en los estudios longitudinales. *Revista de Investigación Educativa*, 3(6), 252-270. Pearson.
- Ding, H. y Homer, M. (2020). Interpreting mathematics performance in PISA: Taking account of reading performance. *International Journal of Educational Research*, 102, 101566. <https://doi.org/10.1016/j.ijer.2020.101566>
- Doncel Abad, D. y Cabrera Álvarez, P. (2020). Comunidades Autónomas bilingües, identidades y desempeño educativo según PISA 2015. *Revista de educación*, 387, 163-188. <https://doi.org/10.4438/1988-592X-RE-2020-387-443>
- Dumay, X. y Dupriez, V. (2014). Educational quasi-markets, school effectiveness and social inequalities. *Journal of Education Policy*, 29(4), 510-531. <https://doi.org/10.1080/02680939.2013.850536>
- Gamazo, A. y Martínez-Abad, F. (2020). An Exploration of Factors Linked to Academic Performance in PISA 2018 Through Data Mining Techniques. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.575167>
- Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., y Rodríguez-Conde, M. J. (2018). Evaluación de factores relacionados con la eficacia escolar en PISA 2015. Un análisis multinivel. *Revista de educación*, 379, 56-84.
- Gaviria Soto, J. L. y Castro Morera, M. (2005). *Modelos jerárquicos lineales*. La Muralla.
- Gómez-Fernández, N. y Mediavilla, M. (2021). Exploring the relationship between Information and Communication Technologies (ICT) and academic performance: A multilevel analysis for Spain. *Socio-Economic Planning Sciences*, 77, 101009. <https://doi.org/10.1016/j.seps.2021.101009>
- Han, S. W. (2018). School-based teacher hiring and achievement inequality: A comparative perspective. *International Journal of Educational Development*, 61, 82-91. <https://doi.org/10.1016/j.ijedudev.2017.12.004>
- Hayes, A. F. (2006). A Primer on Multilevel Modeling. *Human Communication Research*, 32(4), 385-410. <https://doi.org/10.1111/j.1468-2958.2006.00281.x>
- Hu, X., Gong, Y., Lai, C., y Leung, F. K. S. (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Computers & Education*, 125, 1-13. <https://doi.org/10.1016/j.compedu.2018.05.021>
- Joaristi, L., Lizasoain Hernández, L., y Azpillaga, V. (2014). Detección y caracterización de los centros escolares de alta eficacia de la Comunidad Autónoma del País Vasco mediante Modelos Transversales Contextualizados y Modelos Jerárquicos Lineales. *Estudios sobre educación*, 27, 37-61.
- Jornet Meliá, J. M. (2016). Análisis metodológico del proyecto PISA como evaluación internacional. RELIEVE. *Revista Electrónica De Investigación y Evaluación Educativa*, 22(1), 1-26. <https://doi.org/10.7203/relieve22.1.8293>

- Kameshwara, K. K., Sandoval-Hernandez, A., Shields, R., y Dhanda, K. R. (2020). A false promise? Decentralization in education systems across the globe. *International Journal of Educational Research*, 104, 101669. <https://doi.org/10.1016/j.ijer.2020.101669>
- Kyriakides, L., Creemers, B. P. M., y Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12-23. <https://doi.org/10.1016/j.tate.2008.06.001>
- Lafontaine, D., Baye, A., Vieluf, S., y Monseur, C. (2015). Equity in opportunity-to-learn and achievement in reading: A secondary analysis of PISA 2009 data. *Studies in Educational Evaluation*, 47, 1-11. <https://doi.org/10.1016/j.stueduc.2015.05.001>
- Laukaityte, I. y Rolfsman, E. (2020). Low, medium, and high-performing schools in the Nordic countries. Student performance at PISA Mathematics 2003-2012. *Education Inquiry*, 11(3), 276-295. <https://doi.org/10.1080/20004508.2020.1721256>
- Lee, V. E. (2000). Using Hierarchical Linear Modeling to Study Social Contexts: The Case of School Effects. *Educational Psychologist*, 35(2), 125-141. [https://doi.org/10.1207/S15326985EP3502\\_6](https://doi.org/10.1207/S15326985EP3502_6)
- Lenkeit, J. (2012). How effective are educational systems? A value-added approach to study trends in PIRLS. *Journal of Educational Research Online*, 4(2), 143-173. <https://hdl.handle.net/11245/1.405026>
- Li, H. (2016). How is formative assessment related to students' reading achievement? Findings from PISA 2009. *Assessment in Education: Principles, Policy & Practice*, 23(4), 473-494. <https://doi.org/10.1080/0969594X.2016.1139543>
- Martínez-Abad, F. (2019). Identification of Factors Associated With School Effectiveness With Data Mining Techniques: Testing a New Approach. *Frontiers in Psychology*, 10, 2583. <https://doi.org/10.3389/fpsyg.2019.02583>
- Martínez-Abad, F., Gamazo, A., y Rodríguez-Conde, M.-J. (2020). Educational Data Mining: Identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation*, 66, 100875. <https://doi.org/10.1016/j.stueduc.2020.100875>
- Martínez-Abad, F., Lizasoain, L., Castro, M., y Joaristi, L. M. (2017). Selección de escuelas de alta y baja eficacia en Baja California (México). *REDIE: Revista Electrónica de Investigación Educativa*, 19(2), 38-53. <https://doi.org/10.24320/redie.2017.19.2.960>
- Meunier, M. (2011). Immigration and student achievement: Evidence from Switzerland. *Economics of Education Review*, 30(1), 16-38.
- Molina Portillo, E., Contreras García, J. M., Molina Muñoz, D., y Sánchez Pelegrín, J. A. (2022). Estudio por género del impacto de factores contextuales en el rendimiento matemático del alumnado español en PISA 2018. *Revista Complutense de Educación*, 33(4), 645-656. <https://doi.org/10.5209/rced.76428>
- OECD. (2006). *El programa PISA de la OCDE: Qué es y para qué sirve*. <http://iice.institutos.filo.uba.ar/el-programa-pisa-de-la-ocde-qu%C3%A9-es-y-para-qu%C3%A9-sirve>
- OECD. (2019). *PISA 2018 Technical Report* (p. 468).
- Pomianowicz, K. (2021). Educational achievement disparities between second-generation and non-immigrant students: Do school characteristics account for tracking effects? *European Educational Research Journal*, 22(3), 297-324. <https://doi.org/10.1177/147490412111039929>

- Raudenbush, S. W. y Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE.
- Rutkowski, D., Rutkowski, L., Wild, J., y Burroughs, N. (2018). Poverty and educational achievement in the US: A less-biased estimate using PISA 2012 data. *Journal of Children and Poverty*, 24(1), 47-67. <https://doi.org/10.1080/10796126.2017.1401898>
- Scheerens, J., Luyten, H., van den Berg, S. M., y Glas, C. A. W. (2015). Exploration of direct and indirect associations of system-level policy-amenable variables with reading literacy performance. *Educational Research and Evaluation*, 21(1), 15-39. <https://doi.org/10.1080/13803611.2015.1008520>
- Sirin, S. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75. <https://doi.org/10.3102/00346543075003417>
- Snijders, T. A. B. y Bosker, R. J. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE.
- Sullivan, A. y Calderwood, L. (2017). Surveys: longitudinal, cross-sectional, and trend studies. *The BERA/SAGE Handbook of Educational Research: Two Volume Set* (Vol.2, pp.395-415). SAGE Publications Ltd. <https://doi.org/10.4135/9781473983953>
- Willms, J. D. (2010). School Composition and Contextual Effects on Student Outcomes. *Teachers College Record*, 112(4), 1008-1037.
- Wiseman, A. W. (2013). Policy responses to PISA in comparative perspective. En H. D. Meyer y A. Benavot (Eds.), *PISA, power, and policy: The emergence of global educational governance* (pp. 303-322). Symposium Books.
- Wu, H., Shen, J., Zhang, Y., y Zheng, Y. (2020). Examining the effect of principal leadership on student science achievement. *International Journal of Science Education*, 42(6), 1017-1039. <https://doi.org/10.1080/09500693.2020.1747664>
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2), 114-128. <https://doi.org/10.1016/j.stueduc.2005.05.005>
- Yetişir, M. y Bati, K. (2021). The Effect of School and Student-Related Factors on PISA 2015 Science Performances in Turkey. *International Journal of Psychology and Educational Studies*, 8, 170-186. <https://doi.org/10.52380/ijpes.2021.8.2.433>

Fecha de recepción: 15 de noviembre de 2022.

Fecha de revisión: 13 de diciembre de 2022.

Fecha de aceptación: 29 de noviembre de 2023.