

# Impacto en el rendimiento académico de la inmersión lingüística y del índice socioeconómico y cultural

## *The impact of linguistic immersion and economic, social and cultural status on academic performance*

**Francisco LÓPEZ-RUPÉREZ.** Director de la Cátedra de Políticas Educativas. Universidad Camilo José Cela, España ([flopezr@ucjc.edu](mailto:flopezr@ucjc.edu)).

**Álvaro MORALEDA-RUANO.** Coordinador de investigación. Universidad Camilo José Cela, España ([amoraleda@ucjc.edu](mailto:amoraleda@ucjc.edu)).

**Isabel GARCÍA-GARCÍA.** Colaboradora de la Cátedra de Políticas. Universidad Camilo José Cela, España. ([isabel.garcia@ucjc.edu](mailto:isabel.garcia@ucjc.edu)).

### Resumen:

La cuestión lingüística en el ámbito escolar presenta en España una dimensión educativa de notable interés social, toda vez que un proceso inmersivo franco en las lenguas regionales podría estar afectando al rendimiento de los alumnos en función de su extracción lingüística y sociocultural. En este contexto, el presente trabajo tiene como objetivo principal analizar las relaciones entre nivel socioeconómico y cultural (ISEC), porcentaje de alumnos que hablan en la escuela un idioma diferente al de su hogar (idioma) y rendimiento escolar en PISA (rendimiento). En particular, se pretende medir el poder predictivo de la segunda variable con respecto a la tercera desde una perspectiva territorial. Se ha desarrollado una metodología basada en análisis de

regresión lineal bivariada y en modelos de ecuaciones estructurales aplicadas sobre datos de esas tres variables, disponibles en las bases de PISA 2015 y PISA 2022. La realización de estos análisis secundarios ha confirmado: a) la importancia del poder predictivo del ISEC sobre el rendimiento de forma consistente, tanto en 2015 como en 2022; b) la intensificación del vínculo entre ambas variables en dicho período de tiempo; y c) la aparición de una asociación moderada, pero significativa, de la variable idioma como predictor del rendimiento en el ámbito territorial. Se aprecia que el impacto de las diferencias territoriales en materia de ISEC sobre el rendimiento ha aumentado de forma notable entre 2015 y 2022, lo que resulta compatible con el incremento de otras brechas territoriales descritas en la literatura.

---

Fecha de recepción del original: 23-04-2024.

Fecha de aprobación: 05-08-2024.

Cómo citar este artículo: López-Rupérez, F., Moraleda-Ruano, Á., y García-García, I. (2024). Impacto en el rendimiento académico de la inmersión lingüística y del índice socioeconómico y cultural [The impact of linguistic immersion and economic, social and cultural status on academic performance]. *Revista Española de Pedagogía*, 82 (289), 515-537. <https://doi.org/10.22550/2174-0909.4178>

Por otra parte, los modelos de una inmersión lingüística severa marcan diferencias significativas entre territorios en materia de rendimiento que se sustentan, a modo de ejemplo, en un retraso, solo por este motivo, de Cataluña respecto de Madrid de un año escolar aproximadamente.

**Palabras clave:** programas de inmersión, rendimiento académico, equidad educativa, política educativa, evaluación educativa, modelos de ecuaciones estructurales.

### Abstract:

The linguistic issue in the schools presents in Spain an educational dimension of considerable social interest, since a strong immersive process in the regional language could be affecting the performance of students, depending on their linguistic and socio-cultural background. In this context, the main objective of this paper is to analyse the relations between economic, social and cultural status (ESCS); percentage of students who speak a language different from their home (language); and school performance in PISA (performance). And measure the predictive power of the second variable with respect to the third, from

a territorial perspective. A methodology has been developed based on bivariate linear regression analysis and on structural equations modelling, applied on data of these three variables available in PISA 2015 and PISA 2022. These secondary analyses have confirmed: a) the importance of the predictive power of the ESCS on performance in a consistent manner, both in 2015 and 2022; b) the intensification of the association between the two variables in that period; and c) the emergence of a moderate but significant association of the language variable as a predictor of performance at the territorial level. The impact of territorial differences in ESCS on performance has increased significantly between 2015 and 2022, which is compatible with the increase of other territorial divides described in the literature. On the other hand, the models of a strong linguistic immersion mark significant differences between territories, in terms of performance. For example, for this reason alone, Catalonia's delay compared to Madrid is approximately one school year.

**Keywords:** immersion programs, academic achievement, educational equity, educational policy, educational assessment, structural equations modelling.

## 1. Introducción

La cuestión lingüística en el ámbito escolar (entendida, en el caso español, principalmente como un problema complejo asociado a la coexistencia en la escuela de la lengua oficial y, en su caso, de la cooficial de la correspondiente comunidad autónoma) ha ido ganando relevancia conforme los modelos basados en la libertad de elección de las familias (o, en su defecto, en una coexistencia

amable entre las dos lenguas) han virado, de manera progresiva, hacia modelos de inmersión franca en la lengua cooficial. Ese proceso inmersivo podría estar afectando al rendimiento de los alumnos en función de su extracción lingüística o sociocultural.

La Organización de Naciones Unidas para la Educación la Ciencia y la Cultura (UNESCO) hace tiempo que sentó doctrina (UNESCO,

1953) en el sentido de privilegiar el uso de la lengua materna (o primera lengua que el niño aprende) en la educación infantil y primaria. Con posterioridad, con ocasión del Informe de Seguimiento de la Educación en el Mundo (UNESCO, 2016), la organización advertía que «diversas evaluaciones internacionales y regionales del aprendizaje confirman que cuando se hablan lenguas distintas en el hogar y en la escuela, se observa un efecto negativo en los resultados de las pruebas» (p. 24). En un documento más reciente (UNESCO-IIEP, 2021), en el que se aporta una revisión significativa sobre lo que se sabe al respecto (Benson, 2004; Bühmann y Trudell, 2007; Pinnock, 2009a, 2009b), se afirma que

la mayor parte de la investigación ahora ha concluido que los resultados del aprendizaje mejoran cuando a los niños y a las niñas se les enseña en su lengua materna al menos durante los seis primeros años de la escuela primaria, como idioma principal de instrucción, antes de que se introduzca el segundo idioma. (p. 1)

Por otro lado, la investigación internacional ha advertido que la educación basada en la lengua materna beneficia en especial a los sectores desaventajados, incluyendo en este grupo a los del medio rural (Ball, 2011; Hovens, 2002).

En lo que concierne a los países en vías de desarrollo, Reily (2019) alude a una amplia relación de estudios experimentales con grupo de control que se apoyan en pruebas estandarizadas y que demuestran, en una notable diversidad de países, los beneficios del uso de la lengua materna en la escuela. Los documentos institucionales más arriba citados adoptan un posicionamiento que resume

las aportaciones de estudios y revisiones previos alineados con resultados similares. Estos resultados se repiten de un modo consistente en el ámbito de los países desarrollados.

Así, en los Estados Unidos, Thomas y Collier (1997, 2002) realizaron un estudio a gran escala, de carácter longitudinal, basado en los registros de expedientes de cerca de 200 000 alumnos, sobre las disposiciones educativas para estudiantes de lenguas minoritarias en las escuelas públicas norteamericanas y su efecto sobre el rendimiento académico. Encontraron que el predictor más fuerte del rendimiento en la segunda lengua era la cantidad de educación formal en lengua materna.

En el contexto europeo, la posición más destacable es la del Consejo de Europa, que, sin perjuicio de una apuesta clara por el multilingüismo, asume y desarrolla la establecida por la UNESCO en materia de uso de la lengua materna en la escuela. Así se desprende del Documento 10837, de 7 de febrero de 2006, de la Asamblea Parlamentaria que recoge un informe de la Comisión de Cultura, Ciencia y Educación (Parliamentary Assembly, 2006) y en cuyo punto B.12 de su exposición de motivos se posiciona del modo siguiente:

Evidentemente, en las sociedades europeas, el uso cotidiano de la lengua oficial es la principal condición previa para la integración de los niños cuya lengua principal es diferente de la oficial del país o de la región. Sin embargo, una gran cantidad de investigaciones están de acuerdo en un punto: la escolarización inmediata de estos niños en un idioma que no conocen bien o que no conocen en absoluto (aquí, el término

es enseñanza «por inmersión») pone en grave peligro sus posibilidades de éxito académico. Por el contrario, la educación bilingüe basada en la lengua materna es la base del éxito a largo plazo. (p. 2)

Es necesario precisar que los resultados antes descritos no invalidan, en modo alguno, el objetivo deseable (Bialystok *et al.*, 2012) del multilingüismo en la escuela, sino que cuestionan simplemente el uso de algunos procedimientos o estrategias para conseguirlo. En este sentido, Cummins (2001), sobre la base de la investigación, subraya lo siguiente:

El bilingüismo tiene efectos positivos sobre el desarrollo lingüístico y educativo de los niños. Cuando los alumnos continúan desarrollando sus habilidades en dos o más idiomas durante los años de la escuela primaria, adquieren una comprensión más profunda del idioma y cómo usarlo de manera efectiva [...]. La investigación sugiere que los niños bilingües también pueden desarrollar más flexibilidad en su pensamiento como resultado del procesamiento de información a través de dos idiomas diferentes. (p. 17)

No obstante, Cummins (2001) se alinea con las anteriores orientaciones sobre el uso escolar de la lengua materna cuando afirma que «la promoción de la lengua materna en la escuela ayuda a desarrollar no solo la lengua materna sino también las habilidades de los niños en el idioma mayoritario de la escuela» (p. 18). Es lo que él denomina «teoría de la interdependencia».

En el caso específico de España, no abundan los estudios acerca de los efectos de los modelos lingüísticos de enseñanza sobre el alumnado (Clots-Figueras y Masella,

2013; Calero y Choi, 2019; Carabaña y Fernández-Enguita, 2019), en particular en lo relativo al impacto sobre su rendimiento escolar. A este respecto, es de destacar el citado trabajo de Calero y Choi (2019), que está centrado en el rendimiento académico y restringido exclusivamente al caso catalán. Los autores efectúan análisis secundarios a partir de la base de datos de PISA 2015, mediante modelos lineales multinivel, y concluyen la existencia de un efecto negativo de la política de inmersión sobre el rendimiento escolar que sintetizan en los siguientes términos: «Los resultados que hemos expuesto apuntan claramente a la existencia de un problema de equidad, que genera “perdedores” de la política de inmersión lingüística en Cataluña» (p. 16).

Por otra parte, el trabajo de Carabaña y Fernández-Enguita (2019) se centra en la relación entre lenguas regionales en España y el sentimiento de pertenencia al centro de los alumnos. Recurren, para ello, a análisis secundarios sobre la base de datos de PISA 2015 y concluyen que dicho sentimiento (que es relevante tanto para el bienestar de los alumnos como para su rendimiento) es notablemente menor en las CCAA bilingües; más aún, que son «todos los alumnos que estudian en lenguas regionales los que deprimen el nivel de identificación con el centro, independientemente de cuál sea la lengua de su hogar» (p. 1). Los autores plantean como hipótesis explicativa una «diglosia escuela-sociedad»: el sentimiento de pertenencia a la escuela padece cuando la lengua de enseñanza es la de menor preferencia o uso sociales.

Aunque no relacionado directamente con los resultados académicos, el trabajo de Clots-Figueras y Masella (2013) sobre lengua,

educación e identidad muestra, a partir de datos de encuestas realizadas en Cataluña, que los encuestados que han estado expuestos durante más tiempo a la enseñanza en catalán tienen sentimientos catalanes más fuertes. El efecto también parece estar presente entre los individuos cuyos padres no son de origen catalán; además, la reforma afecta, según los autores, a las preferencias políticas y a las actitudes hacia la organización del Estado.

La reciente publicación de los datos de PISA 2022, con una caída notoria de las puntuaciones correspondientes a Cataluña en las tres áreas principales de evaluación (OECD, 2023), ha puesto de nuevo sobre el tapete, tanto en el ámbito social como en el político (López-Rupérez y García-García, 2023; Rodríguez, 2023; Rodríguez-Sanmartín, 2023), la dimensión propiamente educativa de la cuestión lingüística.

No obstante, se trata de una cuestión que, además de a Cataluña, concierne al conjunto de España, por diferentes razones. En primer lugar, porque estamos ante situaciones que afectan a derechos fundamentales cuya garantía corresponde al Estado. Pero, además, porque hasta seis comunidades autónomas (por el momento, Baleares, Cataluña, Comunidad Valenciana, Galicia, Navarra y País Vasco) disponen de modelos de coexistencia de lenguas cooficiales con el español en tanto que lengua oficial en todo el territorio nacional. En algunos casos, dichos modelos lingüísticos han ido aproximándose hacia el de una inmersión franca. Conviene, por tanto, generar evidencias referidas al caso español que permitan reflexionar acerca de la magnitud de los efectos y de sus consecuencias educativas sobre una base empírica.

Por tales motivos, en el presente trabajo, se abordará esta cuestión desde una perspectiva territorial. Se adoptará un enfoque que integre, junto con los aspectos lingüísticos, los socioeconómicos y los de rendimiento académico, y se recurrirá, para ello, a los datos de PISA. Los objetivos mayores de esta investigación son los siguientes:

1. Proporcionar una visión comparativa, mediante un análisis descriptivo, de los resultados de PISA en los años 2015 y 2022, con la presentación detallada de las puntuaciones de rendimiento, ISEC e idioma en diferentes regiones y ciudades autónomas españolas.
2. Construir y evaluar un modelo de ecuaciones estructurales (SEM) sobre la influencia del índice socioeconómico y cultural (ISEC) en el rendimiento académico en ciencias, matemáticas y lectura sobre la muestra de datos de PISA 2015.
3. Analizar el impacto del idioma hablado en casa sobre el rendimiento académico en PISA 2022 mediante un análisis de regresión lineal bivariado.
4. Construir y evaluar un SEM (*structural equation modeling*) sobre la relación entre el ISEC y el porcentaje de alumnos que hablan un idioma diferente al dominante en las escuelas (idioma) y su influencia sobre el rendimiento académico en ciencias, matemáticas y lectura en la muestra de datos de PISA 2022.



## 2. Método

### 2.1. Muestras

En el presente estudio, se han empleado dos conjuntos de datos fundamentales. Por un lado, se ha hecho uso de la muestra nacional PISA (OECD, 2016) correspondiente a la edición de 2015. Esta muestra, compuesta por 6736 estudiantes (39066 cuando se incluyen las muestras ampliadas de las comunidades autónomas), es representativa de la población española de alumnos de 15 años, en su mayoría

inscritos en cuarto curso de la Educación Secundaria Obligatoria (ESO). Por otro, se ha adoptado de forma análoga el muestreo que corresponde a la edición de 2022 de PISA (OECD 2023), que consta de 30800 estudiantes al considerar el total de las muestras disponibles, tanto de las comunidades como de las ciudades autónomas.

La Tabla 1 proporciona información detallada sobre las extensiones de las muestras empleadas en función de los diferentes territorios.

TABLA 1. Tamaño y representatividad de las muestras subnacionales en las ediciones de Pisa 2015 y 2022.

	2015		2022	
	<i>n</i>	%	<i>n</i>	%
España	6736	17.2	-	-
Andalucía	1813	4.6	1707	5.5
Aragón	1798	4.6	1420	4.6
Asturias	1790	4.6	1724	5.6
Baleares	1797	4.6	1576	5.1
Canarias	1842	4.7	1686	5.5
Cantabria	1924	4.9	1677	5.4
Castilla y León	1858	4.8	1733	5.6
Castilla-La Mancha	1889	4.8	1671	5.4
Cataluña	1769	4.5	1597	5.2
Comunidad Valenciana	1625	4.2	1625	5.5
Extremadura	1809	4.6	1685	5.9
Galicia	1865	4.8	1824	4.5
Madrid	1808	4.6	2138	6.9
Murcia	1796	4.6	1669	5.4
Navarra	1874	4.8	1826	5.9
País Vasco	3612	9.2	3244	10.5
La Rioja	1461	3.7	1394	5.3
Ceuta	-	-	345	1.1
Melilla	-	-	259	0.8
Total	39066	100	30800	100.0

Fuente: elaboración propia a partir de las bases de datos de PISA 2015 y de PISA 2022.

## 2.2. Variables

Se han organizado las variables de interés en tres categorías fundamentales y han sido tratadas como variables latentes. En un contexto como el presente, de uso de SEM, es importante señalar que las variables latentes son constructos subyacentes que no son directamente observables, pero se infieren a través de variables observadas. A continuación, se especifican las tres categorías de variables latentes utilizadas en el estudio:

Variable latente rendimiento (rendimiento) desagregada por áreas:

- Rendimiento medio en matemáticas: incluye los valores brutos de la materia de matemáticas evaluados en 2015 (Rto\_Mat\_15) y en 2022 (Rto\_Mat\_22).
- Rendimiento medio en lengua: engloba los valores brutos de la materia de lengua evaluados en 2015 (Rto\_Lec\_15) y en 2022 (Rto\_Lec\_22).
- Rendimiento medio en ciencias: comprende los valores brutos de la materia de ciencias evaluados en 2015 (Rto\_Cnc\_15) y en 2022 (Rto\_Cnc\_22).

Variable latente índice del nivel socioeconómico y cultural (ISEC):

- ISEC de los valores reportados en PISA 2015 (ISEC\_15).
- ISEC de los valores reportados en PISA 2022 (ISEC\_22).

Variable latente porcentaje de alumnos que hablan en casa un idioma diferente al de las escuelas (idioma):

- Porcentaje de alumnos que hablan en casa un idioma diferente al de las escuelas, según los valores del informe PISA de 2022 (Idioma\_22).

## 2.3. Procedimientos de análisis

De conformidad con los objetivos de la investigación, se ha llevado a cabo un análisis que comprende tres enfoques básicos. Como primera aproximación, se ha realizado un análisis comparado de datos relevantes con una orientación eminentemente descriptiva. Para dicha comparación, se han recopilado los valores promedio de las variables descritas con anterioridad correspondientes a los años 2015 y 2022. Como complemento, se ha realizado un análisis de regresión lineal bivariado con la finalidad de evaluar el valor predictivo de la variable idioma con respecto a la variable rendimiento para 2022. Por último, se han empleado diversos métodos estructurales SEM (Bryman y Cramer, 1994; Pérez *et al.*; 2013) para validar un modelo causal que englobe, de manera conjunta, las variables rendimiento e ISEC en 2015; y otro modelo idéntico para 2022, al que se ha añadido, en este caso, los análisis de la variable latente idioma.

El enfoque de ecuaciones estructurales empleado se deriva de una técnica que fusiona el análisis factorial con la regresión lineal. Esta técnica supera la simple medición del impacto directo de las variables independientes sobre la dependiente al considerar la interacción y posible influencia indirecta entre ellas. Aunque no demuestra

causalidad estricta, facilita el razonamiento sobre hipótesis causales que reposan en evidencia empírica (Batista *et al.*, 2000). Para la investigación, se asume el supuesto de que existen, cuanto menos, un orden causal débil y un cierre causal (Kim & Kohout, 1988). Este tipo de aproximación constituye una práctica común en estudios de naturaleza similar (Bulut *et al.*; 2012; Calero y Choi, 2019; López-Rupérez *et al.*, 2021; Salas *et al.*; 2017; Swe *et al.*; 2020).

Para la realización de estos análisis, se ha empleado el *software* de análisis estadístico *Jamovi* (The jamovi project, 2022), en su versión 2.3. Se basa en el paquete de *software* R *lavaan* (Rosseel, 2012) para la modelización de ecuaciones estructurales (Gallucci y Jentschke, 2021) y para la representación gráfica de los diagramas de camino (Epskamp *et al.*, 2019).

### 3. Resultados

#### 3.1. Una visión comparada de los resultados de PISA 2015 y 2022

En la Tabla 2, se muestran los resultados obtenidos en las ediciones 2015 y 2022 de PISA en lo relativo al rendimiento por áreas y promedio, ISEC e idioma. Estos datos proporcionan una visión comparativa, en diferentes regiones, del desempeño académico y de factores asociados, lo que permite aproximarse a una comprensión de las dinámicas educativas y sociales a lo largo de un período suficientemente amplio de tiempo. Además, dicho período obvia de forma deliberada la anomalía en la realización de las pruebas (y, por consiguiente, en sus resultados) que supuso para España la edición de 2018, al igual que ha procedido PISA en los informes por países (OECD, 2023).

TABLA 2. Puntuaciones en rendimiento, ISEC e idioma de las muestras en los niveles OCDE nacional y subnacionales correspondientes a las ediciones PISA 2015 y 2022.

Zona	Rto_Mat		Rto_Lec		Rto_Cnc		Rto_Promedio		ISEC		Idioma
	2015	2022	2015	2022	2015	2022	2015	2022	2015	2022	2022
OCDE	490	472	493	476	493	485	493	478	0	0	11.18
España	486	473	496	474	493	485	491	477	-.51	-.03	19.40
Andalucía	466	457	479	461	473	473	473	464	-.87	-.18	6.31
Aragón	500	487	506	488	508	499	505	491	-.39	.00	10.66
Asturias	492	495	498	497	501	503	497	498	-.42	.09	4.48
Baleares	476	471	485	472	485	480	482	474	-.65	-.05	51.80
País Vasco	492	482	491	466	483	480	489	476	-.25	.18	15.30
Canarias	452	447	483	463	475	473	470	461	-.80	-.21	5.41
Cantabria	495	495	501	494	496	504	497	498	-.43	.03	3.82
Castilla y León	506	499	522	498	519	506	516	501	-.44	.08	3.45



## Impacto en el rendimiento académico de la inmersión lingüística y del índice socioeconómico y cultural

Castilla-La Mancha	486	464	499	468	497	475	494	469	-.66	-.16	6.52
Cataluña	500	469	500	462	504	477	501	470	-.35	-.02	55.54
Com. Valenciana	485	473	499	482	494	483	493	479	-.53	-.10	24.56
Extremadura	473	469	475	468	474	479	474	472	-.79	-.14	2.32
Galicia	494	486	509	485	512	506	505	493	-.52	.07	36.19
La Rioja	505	493	491	487	498	500	498	493	-.46	-.06	8.17
Madrid	503	494	520	496	516	502	513	497	-.01	.25	5.82
Murcia	470	463	486	468	484	482	480	471	-.82	-.24	7.40
Navarra	518	492	514	478	512	489	515	486	-.32	.02	27.75
Ceuta	-	395	-	404	-	410	-	403	-	-.47	26.41
Melilla	-	404	-	405	-	414	-	408	-	-.79	15.92

Fuente: elaboración propia a partir de OECD, 2016 y 2023.

Un análisis sistemático de los datos de la Tabla 2 pone de manifiesto lo siguiente:

- En lo que respecta a las variables de rendimiento, excepción hecha de Asturias, Cantabria y La Rioja, que mejoran, al menos, en una de las áreas, todas las demás empeoran en las tres. Por consiguiente, cuando se toman en consideración las puntuaciones promedio, todas, salvo las dos primeras, empeoran. Cataluña con 31 puntos PISA y Castilla-La Mancha con 25 son las que presentan los mayores retrocesos.
- En lo relativo al ISEC, todas las comunidades autónomas mejoran con respecto a la media OCDE en el período comprendido entre 2015 y 2022, lo que se traduce en que España mejora en .48 puntos y se sitúa por encima de la media actualizada. Destacan en la mejora Andalucía y Baleares, con incrementos

de .69 y .60 puntos del ISEC, respectivamente. Este fenómeno será analizado en el apartado de discusión.

- En lo concerniente al porcentaje de alumnos que hablan en casa un idioma diferente al de la escuela, Cataluña y Baleares, con valores de 55.54% y 51.80%, respectivamente, son las comunidades autónomas con valores máximos para esta variable. En el otro extremo, son Extremadura y Castilla-León, con 2.32% y 3.45%, respectivamente, las que alcanzan los valores mínimos.

### 3.2. Modelo de ecuaciones estructurales correspondiente a PISA 2015

En una primera aproximación, se ha definido un modelo centrado exclusivamente en los datos de 2015 y se ha establecido la relación entre el ISEC y el rendimiento en ciencias, matemáticas y lectura. La aplicación del SEM por medio

del método de máxima verosimilitud con optimización no lineal ha evaluado un modelo que incorpora variables latentes (ISEC y rendimiento) y observadas. Este modelo revela relaciones significativas entre el ISEC\_15 y el rendimiento académico, respaldado por pruebas estadísticas y un ajuste adecuado.

Para evaluar la calidad del ajuste del modelo, se han empleado diversos índices, cuyos resultados se detallan en la Tabla 3; estos índices indican que el modelo pro-

puesto presenta un ajuste razonable, significativamente mejor que un modelo de referencia ( $\chi^2 = 92.48, p < .001$ ). Los índices, tales como la raíz del residuo cuadrático medio estandarizado (SRMR  $< 0.05$ ), la raíz del error cuadrático medio de aproximación (RMSEA  $< 0.05$ ), el índice de bondad de ajuste global corregido (AGFI  $> 0.95$ ), el índice de bondad de ajuste comparativo (CFI  $> 0.90$ ) o el índice Tucker-Lewis (TLI  $> 0.70$ ), respaldan la solidez del modelo, que sigue las recomendaciones de Schermelleh-Engel *et al.* (2003).

TABLA 3. Índices de ajuste del modelo SEM PISA 2015.

	Valores obtenidos
Valor de $p$ en $\chi^2$	.008
$\chi^2/gl$	4.87
GFI	1
AGFI	0.999
SRMR	0.047
RMSEA	0.009
CFI	0.91
TLI	0.731
Índice de ajuste normalizado de Bentler-Bonett (NFI)	0.895
Índice de ajuste de parsimonia (PGFI)	0.143
Índice de ajuste normalizado de parsimonia (PNFI)	0.298
Criterio de Akaike (AIC)	378.019

La Tabla 4 presenta las estimaciones de parámetros, junto con el modelo de medición, para la relación entre el rendimiento académico y el ISEC en el año 2015. En la primera sección, se muestra el coeficiente de camino entre el rendimiento académico y el ISEC. Tal y como se indica, un aumento de 1 unidad en el ISEC se asocia con un aumento estimado de 48.3 unidades en el rendimiento académico, con un intervalo de confianza del 95 % entre 28.1 y 68.4. Este coeficiente es significativo ( $\beta = 0.748, p < .001$ ) y sugiere una relación positiva entre el ISEC y el rendimiento académico.

Se muestran, asimismo, en dicha Tabla 4 las estimaciones para el modelo de medición de las variables latentes. El ISEC\_15,

como variable observada de la variable latente ISEC, tiene un coeficiente estandarizado de 1.000, lo que indica una relación perfecta. De la misma forma, en cuanto al rendimiento académico, se constata que las variables observadas (RTO\_Ciencias\_15, RTO\_Mat\_15, RTO\_Lectura\_15) tienen coeficientes estandarizados de 1.00, 0.99 y 0.90, respectivamente, señal de las relaciones existentes entre estas variables observadas y la variable latente de rendimiento.

Estos coeficientes estandarizados, junto con sus intervalos de confianza y valores de significancia, proporcionan información sobre la calidad y la fuerza de las relaciones modelizadas. En general, la Tabla 4 respalda de modo empírico la validez del modelo propuesto.

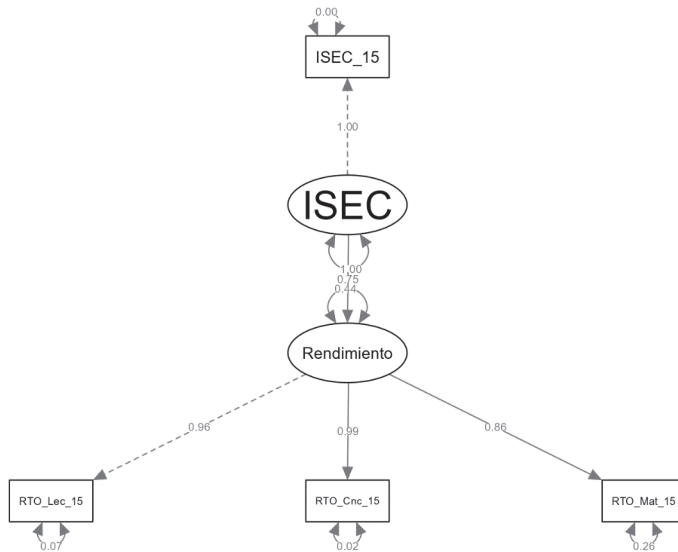
Tabla 4. Estimaciones de parámetros y modelo de medición para la relación entre rendimiento académico, idioma e ISEC en los resultados de PISA 2015.

Variable dependiente	Variable predictora	Estimación	Error estándar	IC 95 % (inf./sup.)		$\beta$	$z$	$p$
Rendimiento	ISEC	48.3	10.3	28.1	68.4	0.748	4.70	<.001
Variable latente	Variable observada							
ISEC	ISEC_15	1.00	0.00	1.00	1.00	1.000		
	RTO_Ciencias_Br_15	1.00	0.00	1.00	1.00	0.992		
Rendimiento	RTO_Mat_Br_15	0.99	0.14	0.71	1.27	0.861	6.84	<.001
	RTO_Lectura_Br_15	0.90	0.07	0.76	1.03	0.965	12.69	<.001

La Figura 1 presenta el *path model* e ilustra el impacto del ISEC en el rendimiento de PISA 2015. El SEM con datos de PISA 2015 subraya la importancia del ISEC en la predicción del rendimiento

académico. En concreto, respalda la hipótesis de que un ISEC más elevado se vincula de forma positiva con un mejor rendimiento en ciencias, matemáticas y lectura.



FIGURA 1. *Path model* de la relación entre ISEC y rendimiento académico en PISA 2015.

### 3.3. El valor predictivo de la variable idioma con respecto al rendimiento en PISA

Para analizar la repercusión de la variable idioma sobre el rendimiento académico en PISA 2022, se llevó a cabo un análisis de regresión lineal bivariado. En la Tabla 5, se presentan métricas clave

que ilustran la conexión entre ambas variables. El coeficiente de correlación múltiple muestra una relación débil, mientras que los bajos valores de  $R^2$  indican que solo una pequeña proporción de la variabilidad del rendimiento se atribuye al idioma.

TABLA 5. Estadísticas de regresión.

Coefficiente de correlación múltiple	0.17
Coefficiente de determinación $R^2$	0.03
$R^2$ ajustado	-0.02
Error típico	26.14
Observaciones	21

En la Tabla 6, se profundiza en la evaluación global del modelo. La falta de significación estadística ( $F = 0.59$ ,  $p > .05$ ) y la modesta suma de cuadrados en la regresión señalan una relación

directa generalmente débil entre el idioma y el rendimiento. Además, los coeficientes asociados a la intercepción y la variable idioma no alcanzan significancia suficiente.

Tabla 6. Análisis de varianza y coeficientes.

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	p
Regresión	1	404.42	404.42	0.59	0.45
Residuos	19	12984.62	683.40		
Total	20	13389.03			

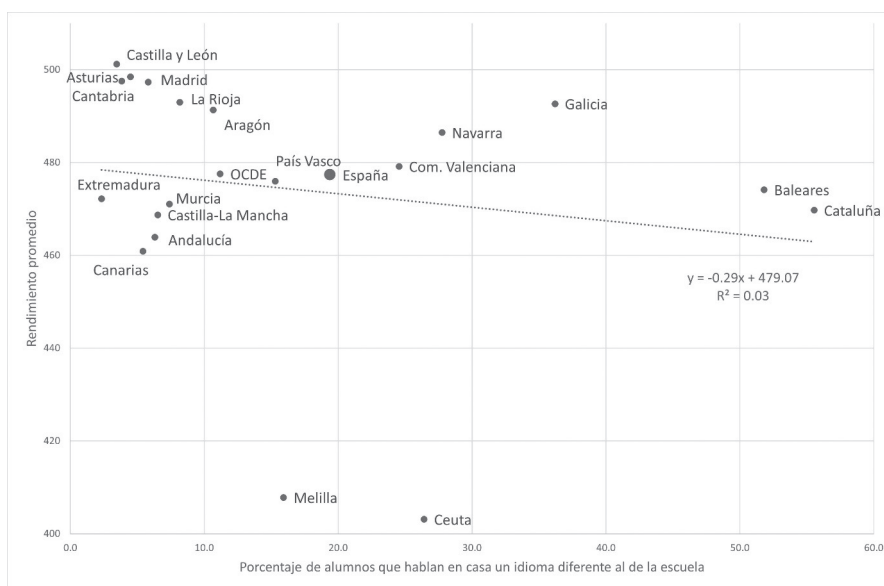
  

	Coefficiente	Error típico	IC 95% (inf./sup.)	t	p
Intercepción	479.07	8.47	461.33 496.81	56.53	0.00
Variable X 1	-0.29	0.38	-1.08 0.50	-0.77	0.45

A pesar de que la variable idioma no muestra un impacto estadísticamente significativo en las puntuaciones de rendimiento en PISA 2022, cuando se analiza de forma gráfica esta débil relación, se

observa (véase la Figura 2) una tendencia de carácter inverso entre ambas variables: rendimiento en PISA 2022 y uso del idioma en muestras en los niveles OCDE, nacional y subnacionales.

FIGURA 2. Análisis de regresión lineal rendimiento PISA 2022 vs. idioma sobre las muestras OCDE, nacional y subnacionales.





### 3.4. Modelo de ecuaciones estructurales correspondiente a PISA 2022

Por último, se ha replicado el enfoque SEM, anteriormente aplicado a 2015, mediante un modelo que opera ahora sobre los resultados de PISA 2022. En este análisis, se establece de nuevo una relación entre el ISEC y el rendimiento en ciencias, matemáticas y lectura, pero se incorporan, además, los resultados relacionados con el porcentaje de alumnos que hablan en casa un idioma diferente al de la escuela (idioma). A pesar de haberse establecido en la sección anterior una relación no significativa, interpretada como ausencia de impacto directo, es importante señalar que un SEM tiene la capacidad de modelar relaciones más complejas entre variables, así como de considerar efectos indirectos y mediaciones.

La aplicación del análisis del SEM mediante el método de máxima verosimilitud con optimización no lineal ha evaluado un modelo que incorpora variables latentes (ISEC, idioma y rendimiento) y observadas. Este modelo revela relaciones significativas entre el ISEC\_22 y el rendimiento académico, así como entre el porcentaje de alumnos que hablan un idioma diferente al de las escuelas (idioma), respaldadas por pruebas estadísticas y un ajuste adecuado.

En relación con el ajuste del modelo, los indicadores que se muestran en la Tabla 7 señalan que el modelo propuesto exhibe un ajuste razonable, considerablemente superior al de un modelo de referencia ( $\chi^2 = 204.6$ ,  $p < .001$ ). Índices tales como SRMR ( $<0.05$ ), RMSEA ( $<0.05$ ), AGFI ( $>0.95$ ), CFI ( $>0.90$ ) y TLI ( $>0.85$ ) respaldan la robustez y validez del modelo.

TABLA 7. Índices de ajuste del modelo SEM en PISA 2022.

	Valores obtenidos
Valor de $p$ en $\chi^2$	.01
$\chi^2/gl$	2.65
GFI	1
AGFI	0.999
SRMR	0.04
RMSEA	0.014
CFI	0.939
TLI	0.869
NFI	0.909
PGFI	0.259
PNFI	0.424
AIC	642.657

La Tabla 8 presenta las estimaciones de parámetros, así como el modelo de medición para la relación entre rendimiento académico, idioma e ISEC en el año 2022. En la sección relacionada con el rendimiento académico, se observa que un incremento de 1 unidad en el ISEC se asocia con un aumento estimado de 166.78 unidades en el rendimiento académico, con un intervalo de confianza del 95% entre 98.78 y 234.78. Este coeficiente es significativo ( $p < .001$ ) e indica

una relación positiva entre ISEC y rendimiento académico. Asimismo, para el predictor idioma, se encuentra que un aumento de 1 unidad en el porcentaje de alumnos que hablan un idioma diferente al de las escuelas se relaciona con una disminución estimada de 0.43 unidades en el rendimiento académico, con un intervalo de confianza del 95% entre -0.81 y -0.04. Este coeficiente también es significativo ( $p < .05$ ) y señala una relación negativa entre idioma y rendimiento.

Tabla 8. Estimaciones de parámetros y modelo de medición para la relación entre rendimiento académico, idioma e ISEC en los resultados de PISA 2022.

Variable dependiente	Variable predictora	Estimación	Error estándar	IC 95% (inf./sup.)		$\beta$	$z$	$p$
Rendimiento	ISEC	166.78	34.69	98.78	234.78	1.037	4.81	<.001
Rendimiento	Idioma	-0.43	0.20	-0.81	-0.04	-0.245	-2.16	<.05

Variable latente	Variable observada	Estimación	Error estándar	IC 95% (inf./sup.)		$\beta$	$z$	$p$
ISEC	ISEC_no_inmig_22	1.00	0.00	1.00	1.00	0.805		
	ISEC_inmig_22	1.53	0.38	0.78	2.28	0.723	3.99	<.001
Idioma	Porc_idiomas_22	1.00	0.00	1.00	1.00	1.000		
	RTO_Mat_22	1.00	0.00	1.00	1.00	0.977		
Rendimiento	RTO_Lengua_22	0.92	0.05	0.82	1.02	0.992	18.18	<.001
	RTO_Ciencias_22	0.94	0.05	0.85	1.05	0.995	19.09	<.001

En lo que concierne a las variables latentes y observadas, cabe destacar que las relaciones entre ISEC y sus componentes, así como entre idioma y el porcentaje de alumnos que hablan diferentes idiomas, son significativas y coherentes con el modelo propuesto. Las variables observadas de rendimiento en las asignaturas

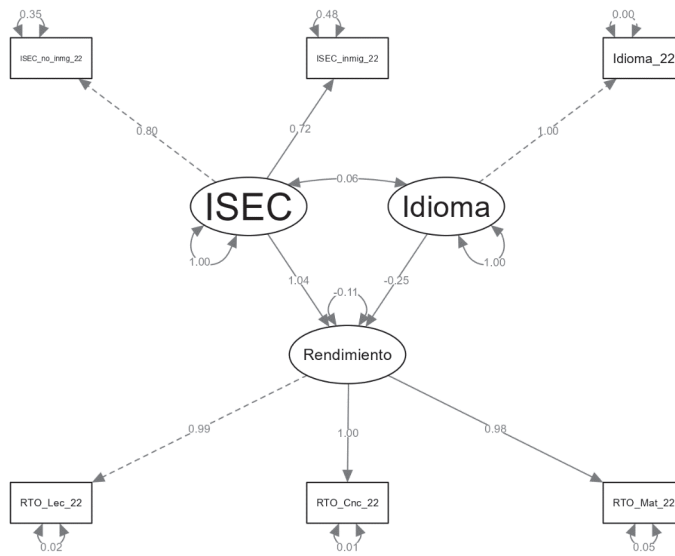
específicas también presentan relaciones positivas y significativas con la variable latente de rendimiento. En conjunto, estas estimaciones detalladas ofrecen una comprensión más profunda de las interacciones entre las variables clave en el contexto del modelo SEM aplicado para el año 2022.



En resumen, en cuanto a las relaciones específicas de PISA 2022, se ha observado una asociación positiva y fuerte entre ISEC y rendimiento, así como una relación negativa de intensidad moderada, pero significativa, entre idioma y rendimiento académico. Estas conclusiones son visualmente respaldadas por el *path model* que se representa en la Figura 3.

En resumen, en cuanto a las relaciones específicas de PISA 2022, se ha observado una asociación positiva y fuerte entre idioma y rendimiento académico. Estas conclusiones son visualmente respaldadas por el *path model* que se representa en la Figura 3.

FIGURA 3. *Path model* de la relación entre ISEC e idioma con el rendimiento académico en PISA 2022.



## 4. Discusión

Como cuestión preliminar, procede aquí subrayar que el presente trabajo se alinea con la revalorización de los análisis de camino [postulada por Judea Pearl (Pearl & Mackenzie, 2020)] como procedimiento que va más allá de un mero ejercicio de estadística, aunque lo contenga. Se basa en una comprensión del problema por parte de cada científico que se refleja en un diagrama causal y que es sometido después a la prueba de la coordinación con la realidad, medida por los datos y por sus análisis cuantitativos (pp. 99-100).

### 4.1. Sobre la influencia del ISEC en el rendimiento

A la hora de analizar la evolución del ISEC entre 2015 y 2022 e interpretar los avances experimentados por todas las comunidades autónomas y por el conjunto de España, es preciso tomar en consideración el hecho de que el ISEC es un indicador normalizado (establecido sobre el conjunto de los países miembros de la OCDE) con un valor medio igual a 0 y una desviación típica igual a 1 (OECD, 2023). En ese período de tiempo, se han incorporado cuatro países de inferior nivel de desarrollo

[Letonia (2016), Lituania y Colombia (2018), y Costa Rica (2021)], lo que explicaría, al menos en parte, el hecho de que España haya mejorado, en términos relativos a la media, el valor de su ISEC.

Por otro lado, y dado el valor predictivo que posee el ISEC sobre el rendimiento, ello tendría que haberse traducido en la correspondiente mejora de los resultados con respecto a la media de la OCDE, lo que no se observa (véase la Tabla 2). Esto indica, de forma plausible, que otros factores de signo contrario han estado operando en dicho período.

Los resultados derivados del primer modelo, a pesar de apoyarse en conglomerados de datos agrupados por comunidades autónomas, son coherentes con la relación establecida de manera empírica en las sucesivas ediciones de PISA que confirma el valor predictivo del ISEC con relación al rendimiento. Así, por ejemplo, el valor del coeficiente de correlación entre esas dos variables definido en PISA 2015 para el conjunto de los países participantes alcanzó la cifra de 0.129 (OECD, 2016). Un análisis de regresión lineal propio, circunscrito a España en el área de ciencias, aportó un valor de  $R = 0.38$  (López-Rupérez *et al.*, 2018, 2019a).

De acuerdo con el informe final de PISA 2022 (OECD, 2023), 20 puntos PISA representan el «ritmo anual de aprendizaje», de modo que las cifras anteriores revelan, a la postre, retrasos estimados del orden de dos años por cada punto del ISEC (de acuerdo con la Tabla 4, una unidad en el ISEC se asocia con un aumento estima-

do de 48.3 unidades PISA en rendimiento académico). Aun cuando el dato del ritmo de desfase que aporta la OCDE es el promedio de los países participantes, facilita una base para efectuar una estimación aproximada de lo que predicen diferencias observadas en cuanto al ISEC entre comunidades autónomas (véase la Tabla 2). Así, por ejemplo, las diferencias entre Andalucía (-0.86) y la Comunidad de Madrid (-0.01) comportarían, de acuerdo con los resultados más arriba descritos, desfases entre ambas regiones equivalentes a cerca de dos años escolares. Ello señala la necesidad de políticas de Estado en materia de equidad interterritorial.

#### 4.2. Sobre los resultados del uso del SEM (*structural equation modeling*)

El análisis del valor predictivo del porcentaje de alumnos que hablan en casa un idioma diferente al de las escuelas con respecto a las puntuaciones promedio en PISA 2022 ha aportado resultados insuficientes. Por tanto, si no se consideran otras variables intervinientes, la variable idioma por sí sola posee un escaso valor predictivo con respecto al rendimiento en PISA. No obstante, es crucial señalar que análisis más detallados, como los realizados mediante modelos SEM, han modificado esta primera conclusión y han permitido una comprensión más completa de las relaciones subyacentes.

La aplicación del modelo SEM sobre PISA 2022 no solo replica las relaciones identificadas en 2015, sino que también amplía la comprensión al incorporar la variable idioma; así, enriquece el análisis al explorar cómo las dinámicas lingüísticas

pueden influir en el rendimiento académico. Esta inclusión fortalece la base empírica en comparación con el modelo de 2015. También introduce la variable tiempo de forma implícita al considerar la situación siete años después. De acuerdo con los resultados mencionados con anterioridad, el impacto de las diferencias entre comunidades autónomas en términos de nivel socioeconómico y cultural medido por el ISEC en 2022 se ha multiplicado por 3.5 en comparación con 2015. En términos relativos, esto sugiere una pérdida de equidad interterritorial en nuestro sistema educativo.

Esta conclusión es compatible con la evolución en las brechas territoriales observadas en otros estudios para variables de resultados, similares y diferentes. Así, por ejemplo, la brecha relativa al parámetro de impacto (medido por la pendiente de la recta de regresión rendimiento *vs.* ISEC) ha aumentado de 11.3 a 14 puntos entre 2015 y 2022 (López-Rupérez *et al.*, 2018 y 2019b; López-Rupérez y García-García, 2023). Aunque referido a variables diferentes, un estudio de Soler (2022) efectuado sobre el período 2005-2021 en materia de abandono educativo temprano y con dos metodologías distintas de análisis ha revelado, de un modo consistente, «la ausencia de convergencia entre las distintas regiones españolas en el periodo analizado, sobre todo en el caso de los hombres» (p. 153).

No obstante esta tendencia, investigaciones previas y los propios análisis de la OCDE (OECD, 2023; Sanz & Tena, 2023) han señalado el efecto parcial de la pandemia del covid-19 en la caída de las puntuaciones de PISA 2022. Asimismo, es posible

conjeturar, sobre la base de evidencias disponibles, que ese efecto no solo se ha dejado sentir sobre las puntuaciones medias, sino también, y de forma preferente, sobre los sectores en desventaja, lo que explicaría el notable aumento del impacto del ISEC sobre el rendimiento descrito en esta investigación. Así, el estudio de Woessmann *et al.* (2020) ha concluido que los alumnos procedentes de entornos desaventajados en términos socioeconómicos han sido los más afectados por la pandemia. Chetty *et al.* (2020) han establecido una reducción del 60% en los avances en rendimiento académico de los estudiantes del cuartil inferior de ingresos. Y Bacher-Hick *et al.* (2021) han aportado una evidencia que podría explicar los resultados anteriores, y es que los crecimientos más altos en el uso de los recursos de la educación *online* se producen en las zonas urbanas y de ingresos más elevados.

Además de las relaciones estadísticamente significativas entre el ISEC<sub>22</sub> y el rendimiento académico, el modelo revela la conexión entre este y el porcentaje de alumnos que hablan en casa un idioma diferente al de las escuelas (idioma). Estas asociaciones respaldan la idea de que tanto el contexto socioeconómico como las variables lingüísticas contribuyen, de manera significativa, al rendimiento estudiantil.

En lo que concierne al predictor idioma en el seno del correspondiente modelo SEM, se ha encontrado que un aumento de 1 unidad en el porcentaje de alumnos que hablan un idioma diferente al de las escuelas se relaciona con una disminución estimada de 0.426 unidades en el rendimiento



académico, con una asociación moderada, pero negativa, entre idioma y rendimiento que es estadísticamente significativa. Ello comporta que el porcentaje de alumnos que hablan en la escuela un idioma distinto del que emplean en casa marca diferencias en el rendimiento académico medido mediante las pruebas de PISA y contribuye, así, a generar brechas de rendimiento entre comunidades autónomas. Aunque, con toda probabilidad, otros factores estarán mediando en la relación entre estas dos variables, no cabe duda de que, en este orden de ideas, los modelos lingüísticos de inmersión severa se asocian a una diferencia significativa entre los territorios en materia de rendimiento cuando se considera el nivel socioeconómico y cultural. Por ejemplo, la diferencia en dichos porcentajes entre Cataluña y Madrid es de, aproximadamente, 50 puntos porcentuales (véase la Tabla 2). Esto, de acuerdo con el modelo SEM, se corresponde con 21 puntos PISA de rendimiento escolar y comporta un retraso académico, solo por ese motivo (de acuerdo con la Tabla 2, la diferencia total es de 27 puntos PISA), de algo más de un año académico entre esas dos comunidades autónomas. Ello apunta a la conveniencia de revisar los modelos de inmersión lingüística en la escuela tomando en consideración su impacto sobre el rendimiento académico.

### 5. Conclusiones

A pesar de la sofisticación del modelo SEM utilizado, la correlación identificada entre las variables no implica de modo necesario causalidad directa. Estas limitaciones abren la puerta a futuras investi-

gaciones que aborden dichas restricciones metodológicas. Por otra parte, aunque el modelo analiza variables clave como ISEC, idioma y rendimiento, se han omitido otros factores potencialmente influyentes, como la calidad del profesorado (Castro y Egido, 2024), la calidad de la dirección escolar (Gairín, 2024) u otros empíricamente relevantes (Hattie, 2023) que podrían estar incidiendo, junto con las variables anteriores, sobre el rendimiento académico. Por tal motivo, la complicación del modelo con la incorporación de variables adicionales constituiría una dirección fructífera para la investigación futura.

Esa cautela imprescindible, junto con la anterior recomendación prospectiva, no son óbice para señalar la coherencia de los resultados obtenidos en la presente investigación con otros ya consolidados en la literatura especializada, algunos de los cuales se han citado más arriba.

Así, los modelos empleados tanto para la edición de PISA de 2015 como para la más reciente de 2022 sugieren que un incremento en el ISEC se asocia con un mejor desempeño académico en el momento de la evaluación, dado que se ha encontrado, de manera consistente, una relación significativa entre un ISEC más elevado y un rendimiento académico superior. Ello concuerda con la asociación reiterada y positiva entre el nivel socioeconómico y cultural de los alumnos y su rendimiento académico que las diferentes ediciones de PISA han puesto de manifiesto de manera reiterada (OECD, 2016, 2023), al igual que otras investigaciones referidas a España (López-Rupérez *et al.*, 2019c).

Cuando se aplica un modelo que toma en consideración, junto con las variables idioma y rendimiento, la variable ISEC, entonces se advierte que el hecho de hablar en la escuela un idioma diferente al empleado en casa puede tener, en los modelos de inmersión franca, un impacto significativo y negativo sobre los resultados académicos de PISA 2022. Ello resulta consistente con esa gran cantidad de investigaciones que han confirmado que la educación basada en la lengua materna mejora los resultados académicos y aumenta de forma significativa las posibilidades del éxito escolar, tal y como ha sido subrayado, entre otras instituciones, por la UNESCO (2021) y el Consejo de Europa (Parliamentary Assembly, 2006).

En definitiva y más allá de argumentos que se pueden invocar al margen de la educación, la aplicación de ese principio de precaución [formulado según el *dictum* clásico *Primum non nocere* ('Lo primero no hacer daño') y que constituye una exigencia deontológica no solo en el ámbito de la salud, sino también en el de la educación] apunta a la necesidad de alinear las políticas educativas y sus resultados con las exigencias de la equidad y del interés superior del menor.

## Contribuciones de los autores

**Francisco López-Rupérez:** Administración del proyecto; Conceptualización; Escritura (borrador original); Supervisión.

**Álvaro Moraleda-Ruano:** Conceptualización; Curación de datos; Escritura (borrador original); Visualización.

**Isabel García-García:** Curación de datos; Visualización.

## Referencias bibliográficas

- Bacher-Hicks, A., Goodman, J., y Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time [Desigualdad en la adaptación de los hogares a las crisis escolares: participación en el aprendizaje en línea inducida por covid en tiempo real]. *Journal of Public Economics*, 193, 104345. <https://doi.org/10.1016/j.jpubeco.2020.104345>
- Ball, J. (2011). *Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years [Mejorar el aprendizaje de los niños de diversos orígenes lingüísticos: educación bilingüe o plurilingüe basada en la lengua materna en los primeros años]*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000212270>
- Benson, C. (2004). *The importance of mother tongue-based schooling for educational quality [La importancia de la escolarización basada en la lengua materna para la calidad de la educación]*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000146632>
- Bialystok, E., Craik, F. I., y Luk, G. (2012). Bilingualism: Consequences for mind and brain [Bilingüismo: consecuencias para la mente y el cerebro]. *Trends in Cognitive Sciences*, 16 (4), 240-250.
- Bryman, A., y Cramer, D. (1994). *Quantitative data analysis for social scientists [Análisis cuantitativo de datos para científicos sociales]*. Routledge.
- Bühmann, D., y Trudell, B. (2007). *Mother tongue matters: Local language as a key to effective learning [La lengua materna importa: la lengua local como clave para un aprendizaje eficaz]*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000161121>
- Bulut, O., Delen, E., y Kaya, F. (2012). An SEM model based on PISA 2009 in Turkey: How does the use of technology and self-regulation activities predict reading scores? [Un modelo SEM basado en PISA 2009 en Turquía: ¿cómo predice

- el uso de la tecnología y las actividades de autorregulación las puntuaciones de lectura?]. *Proceedia - Social and Behavioral Sciences*, 64, 564-573. <https://doi.org/10.1016/j.sbspro.2012.11.066>
- Calero, J., y Choi, Á. (2019). *Efectos de la inmersión lingüística sobre el alumnado castellanoparlante en Cataluña*. Fundación Europea Sociedad y Educación. <https://sociedad.yeducacion.org/site/wp-content/uploads/SE-Inmersion-Cataluna.pdf>
- Carabaña, J., y Fernández-Enguita, M. (2019). Enseñanza en lenguas regionales y sentimiento de pertenencia a la escuela. *Journal of Supranational Policies of Education*, (9), 57-94. <http://doi.org/10.15366/jospoe2019.9.002>
- Castro, M., y Egidio, I. (2024). *Qué sabemos sobre el profesorado. Políticas, evidencias y perspectivas de futuro*. Narcea.
- Chetty, R., Friedman, J. N., Hendren, N., y Stepner, M. (2020). How did covid-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data [¿Cómo afectaron el covid-19 y las políticas de estabilización al gasto y al empleo? Un nuevo rastreador económico en tiempo real basado en datos del sector privado]. *NBER Working Paper*, (27431). <https://doi.org/10.3386/w27431>
- Clots-Figueras, I., y Masella, P. (2013). Education, language and identity [Educación, lenguaje e identidad]. *The Economic Journal*, 123 (570), 332-357.
- Cummins, J. (2001). Bilingual children's mother tongue: Why is it important for education? [Lengua materna de los niños bilingües: ¿Por qué es importante para la educación?]. *Sprogforum*, 7 (19), 15-20.
- Epskamp S., Stuber S., Nak J., Veenman M., y Jorgensen T.D. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output [semPlot: diagramas de camino y análisis visual de la salida de varios paquetes SEM]* [Software]. <https://CRAN.R-project.org/package=semPlot>
- Foguet, J. M. B., y Gallart, G. C. (2000). *Modelos de ecuaciones estructurales*. La Muralla.
- Gairín, J. (2024). *Dirección y liderazgo de los centros educativos. Naturaleza, desarrollo y práctica profesional*. Narcea.
- Gallucci, M., y Jentschke, S. (2021). *SEMLj: jamovi SEM Analysis [Software]*. <https://semlj.github.io/>
- Hattie, J. (2023). *Visible learning: The sequel. A synthesis of over 2,100 meta-analyses relating to achievement [Aprendizaje visible: la secuela. Una síntesis de más de 2100 metaanálisis relacionados con el rendimiento]*. Routledge.
- Hovens, M. (2002). Bilingual education in West Africa: Does it work? [Educación bilingüe en África Occidental: ¿funciona?]. *International Journal of Bilingual Education and Bilingualism*, 5 (5), 249-266.
- Kim, J. O., y Kohout, F. J. (1988). Special topics in general linear models [Temas especiales en modelos lineales generales]. En N. H. Nie, D. H. Bent, y C. H. Hull, *Statistical package for the social sciences [Paquete estadístico para las ciencias sociales]*. McGraw- Hill.
- López-Rupérez, F., García-García, I., y Expósito Casas E. (2018). *PISA 2015 y las comunidades autónomas españolas. Diagnósticos y políticas de mejora*. Universidad Camilo José Cela. [https://www.ujc.edu/pdfs/universidad/pisa2015\\_ccaa.pdf](https://www.ujc.edu/pdfs/universidad/pisa2015_ccaa.pdf)
- López-Rupérez, F., García-García, I., y Expósito-Casas, E. (2019a). Rendimiento en ciencias, concepciones epistémicas y vocaciones STEM en las comunidades autónomas españolas. Evidencias desde PISA 2015, políticas y prácticas de mejora. *Revista Española de Pedagogía*, 272 (77), 5-27. <https://doi.org/10.22550/REP77-1-2019-09>
- López-Rupérez, F., García-García, I., y Expósito-Casas, E. (2019b). Educational effectiveness, efficiency, and equity in Spanish regions: What does PISA 2015 reveal? [Eficacia, eficiencia y equidad educativas en las regiones españolas: ¿Qué revela PISA 2015?]. *ORBIS SCHOLAE*, 12 (2), 9-36. <https://doi.org/10.14712/23363177.2018.291>
- López-Rupérez, F., García-García, I., y Expósito-Casas, E. (2019c). Igualdad de oportunidades e inclusión educativa en España. *RELIEVE*, 25 (2), 1. <http://doi.org/10.7203/relieve.25.2.14351>
- López-Rupérez, F., García-García, I., y Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España: análisis empíricos y recomendaciones políticas. *Revista de Educación*, 394, 325-354. <https://doi.org/10.4438/1988-592X-RE-2021-394-510>
- López-Rupérez, F., y García-García, I. (2023, 11 de diciembre). PISA 2022 y la cuestión territorial en España. *Magisnet*. <https://www.magisnet.com/2023/12/pisa-2022-y-la-cuestion-territorial-en-espana/>
- OECD (2016). *PISA 2015 results (volume I). Excellence and equity in education [Resultados PISA 2015 (volumen I). Excelencia y equidad*

- en la educación]. OECD Publishing. <https://doi.org/10.1787/9789264266490-en>
- OECD (2023). *PISA 2022 results (volume I). The state of learning and equity in education [Resultados PISA 2022 (volumen I). El estado del aprendizaje y la equidad en la educación]*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Parliamentary Assembly. (2006). *The place of mother tongue in school education [El lugar de la lengua materna en la educación escolar]* (Doc. 10837). Committee on Culture, Science and Education. Council of Europe. <https://assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewHTML.asp?FileID=11142&lang=EN>
- Pearl, J., y Mackenzie, D. (2020). *El libro del porqué. La nueva ciencia de la causa y el efecto*. Pasado y Presente.
- Pérez, E., Medrano, L. A., y Sánchez-Rosas, J. (2013). El *path analysis*: conceptos básicos y ejemplos de aplicación. *Revista Argentina de Ciencias del Comportamiento*, 5 (1), 52-66.
- Pinnock, H. (2009a). *Language and education. The missing link. How the language used in schools threatens the achievement of education for all [Lenguaje y educación. El eslabón perdido. Cómo el lenguaje utilizado en las escuelas amenaza el logro de la educación para todos]*. CfBT Education Trust & Save the Children. <https://resource-centre.savethechildren.net/pdf/1674.pdf>
- Pinnock, H. (2009b). *Steps towards learning: A guide to overcoming language barriers in children's education [Pasos hacia el aprendizaje: una guía para superar las barreras lingüísticas en la educación de los niños]*. Save the Children. <https://www.savethechildren.org.uk/content/dam/global/reports/steps-towards-learning-lr.pdf>
- Reilly, C. (2019). *Language and development: Issues in mother tongue based multilingual education [Lenguaje y desarrollo: problemas en la educación multilingüe basada en la lengua materna]* [Documento Inédito]. [https://www.researchgate.net/publication/343768795\\_Language\\_and\\_development\\_Issues\\_in\\_mother-tongue\\_based\\_multilingual\\_education#fullTextFileContent](https://www.researchgate.net/publication/343768795_Language_and_development_Issues_in_mother-tongue_based_multilingual_education#fullTextFileContent)
- Rodríguez, P. (2023, 7 de diciembre). Catalunya se pregunta por qué se ha hundido más que el resto de España en PISA. *Eldiario.es*. [https://www.eldiario.es/catalunya/catalunya-pregunta-hundido-resto-espana-pisa\\_1\\_10748162.html](https://www.eldiario.es/catalunya/catalunya-pregunta-hundido-resto-espana-pisa_1_10748162.html)
- Rodríguez-Sanmartín, O. (2023, 19 de diciembre). Los alumnos de Cataluña son los que muestran el mayor rechazo a la escuela de toda España. *El Mundo*. <https://www.elmundo.es/espana/2023/12/19/6581fb22fdddf570d8b45b5.html>
- Rosseel, Y. (2012). Ijavaan: An R package for structural equation modeling [Ijavaan: un paquete de R para la modelización de ecuaciones estructurales]. *Journal of Statistical Software*, 48 (2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Salas, A. A., Navarro, R., y Montero, E. (2017). Un modelo de ecuaciones estructurales para el estudio de factores que afectan la competencia lectora y la alfabetización matemática: una aproximación bayesiana con datos de PISA 2009. *Estadística Española*, 59 (194), 167-192.
- Sanz, I., y Tena, J.D. (2023). The impact of the covid-19 pandemic on education learning [El impacto de la pandemia de covid-19 en el aprendizaje educativo]. En J. Sanz, e I. Sanz (Eds.), *Addressing inequities in modern educational assessment [Hacer frente a las desigualdades en la evaluación educativa moderna]* (pp. 15-36). Springer. [https://doi.org/10.1007/978-3-031-45802-6\\_2](https://doi.org/10.1007/978-3-031-45802-6_2)
- Schermelleh-Engel, K., Moosbrugger, H., y Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures [Evaluación del ajuste de los modelos de ecuaciones estructurales: pruebas de significación y medidas descriptivas de bondad de ajuste]. *Methods of Psychological Research*, 8 (2), 23-74.
- Soler, A. (2022). Heterogeneidad regional en el abandono educativo temprano. *Magisterio*, 24-25.
- Swe, K.M., Fraser, B. J., y Afari, E. (2020). Structural relationships between learning environments and students' non-cognitive outcomes: Secondary analysis of PISA data [Relaciones estructurales entre los entornos de aprendizaje y los resultados no cognitivos de los estudiantes: análisis secundario de los datos de PISA]. *Learning Environments Research*, 23 (3), 395-412. <https://doi.org/10.1007/s10984-020-09313-2>
- The jamovi project. (2022). *jamovi* (version 2.3) [Software]. <https://www.jamovi.org>
- Thomas, W. P., y Collier, V. P. (1997). *School effectiveness for language minority students [Eficacia escolar para estudiantes de minorías lingüísticas]*



tics]. National Clearinghouse for Bilingual Education. [https://ncela.ed.gov/sites/default/files/legacy/files/rcd/BE020890/School\\_effectiveness\\_for\\_langu.pdf](https://ncela.ed.gov/sites/default/files/legacy/files/rcd/BE020890/School_effectiveness_for_langu.pdf)

Thomas, W. P., y Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement [Un estudio nacional sobre la eficacia de las escuelas para el logro académico a largo plazo de los estudiantes de minorías lingüísticas]*. Center for Research on Education, Diversity and Excellence. <https://escholarship.org/uc/item/65j213pt>

UNESCO. (1953). *The use of the vernacular languages in education [El uso de lenguas vernáculas en educación]*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO).

UNESCO (2016). *Si no entiendes, ¿cómo puedes aprender? Informe de seguimiento de la educación en el mundo*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). [https://unesdoc.unesco.org/ark:/48223/pf0000243713\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000243713_spa)

UNESCO-IIEP (2021, 23 de septiembre). *La lengua de enseñanza*. <https://bit.ly/3UsQE2C>

Woessmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., y Zierow, L. (2020). Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen? [La educación en la crisis del coronavirus: ¿Cómo pasaron los escolares el tiempo que las escuelas estuvieron cerradas y qué medidas educativas defienden los alemanes?]. *ifo Schnelldienst*, 9, 25-39.

## Biografía de los autores

**Francisco López-Rupérez.** Doctor en Ciencias Físicas por la Universidad Complutense de Madrid y Catedrático de Física y Química de Instituto de Educación Secundaria. Ha sido director general de centros educativos del Ministerio de Educación, Cultura y Deporte; secretario general de Educación y Formación Profesional de dicho Ministerio, y presidente del Consejo Escolar del Estado. Es autor

de una docena de libros y de un centenar de artículos en revistas españolas y extranjeras. En la actualidad, es director de la Cátedra Extraordinaria de Políticas Educativas de la Universidad Camilo José Cela y del Máster Universitario en Políticas y Gobernanza de los Sistemas Educativos.

 <https://orcid.org/0000-0003-2613-9652>

**Álvaro Moraleda-Ruano.** Doctor en Pedagogía por la Universidad Complutense de Madrid, con premio extraordinario de doctorado. Su contribución al Departamento de Teoría e Historia de la Educación se refleja en varias becas de colaboración, incluyendo una beca honorífica. Ha enfocado su labor en áreas clave como la inteligencia emocional y los métodos de investigación estadística educativa. En la actualidad, ocupa el cargo de coordinador de investigación en la Facultad de Educación de la Universidad Camilo José Cela. Cuenta con una amplia trayectoria de publicaciones científicas y ha participado en varios grupos y proyectos de investigación financiados externamente.

 <https://orcid.org/0000-0002-3638-8436>

**Isabel García-García.** Licenciada en Ciencias Matemáticas por la Universidad Complutense de Madrid y Catedrática de Matemáticas de Instituto de Educación Secundaria. Ha sido jefa de área del Consejo Escolar del Estado y responsable del *Informe sobre el estado del sistema educativo* en sus ediciones de 2012 a 2016. En la actualidad, es colaboradora de la Cátedra de Políticas Educativas de la Universidad Camilo José Cela.

 <https://orcid.org/0000-0002-9884-6148>



# The impact of linguistic immersion and economic, social and cultural status on academic performance

## Impacto en el rendimiento académico de la inmersión lingüística y del índice socioeconómico y cultural

**Francisco LÓPEZ-RUPÉREZ.** Director of the Chair in Educational Policies. Universidad Camilo José Cela, Spain ([flopezr@ucjc.edu](mailto:flopezr@ucjc.edu)).

**Álvaro MORALEDA-RUANO.** Research coordinator. Universidad Camilo José Cela, Spain ([amoraleda@ucjc.edu](mailto:amoraleda@ucjc.edu)).

**Isabel GARCÍA-GARCÍA.** Collaborator with the Chair in Educational Policies. Universidad Camilo José Cela, Spain ([isabel.ggarcia@ucjc.edu](mailto:isabel.ggarcia@ucjc.edu)).

### Abstract:

The question of language in school has an educational dimension of considerable social interest in Spain since a strong process of immersion in regional languages might be affecting students' performance depending on their linguistic and socio-cultural background. In this context, the paper's aim is to analyse the relations between economic, social, and cultural status (ESCS); the percentage of students who speak a different language at school than at home (language) and school performance in PISA (performance); and in particular to measure the predictive power of the second variable with respect to the third, from a territorial perspective. A methodology was developed based on bivariate linear regression

analysis and structural equation modelling, and was applied to data regarding these three variables from the PISA 2015 and PISA 2022 databases. These secondary analyses have confirmed: (a) the importance of the predictive power of ESCS on performance consistently in both 2015 and 2022; (b) the intensification of the association between the two variables in that period; and (c) the emergence of a moderate but significant association of the language variable as a predictor of Performance at the territorial level. The impact of regional ESCS differences on performance increased significantly between 2015 and 2022, which is compatible with an increase in other territorial divides described in the literature. Moreover, the strong linguistic immersion models display

---

Date of receipt of the original: 2024-04-23.

Date of approval: 2024-08-05.

Please, cite this article as follows: López-Rupérez, F., Moraleda-Ruano, Á., & García-García, I. (2024). The impact of linguistic immersion and economic, social and cultural status on academic performance [Impacto en el rendimiento académico de la inmersión lingüística y del índice socioeconómico y cultural]. *Revista Española de Pedagogía*, 82 (289), 515-537. <https://doi.org/10.22550/2174-0909.4178>

significant differences between territories in terms of performance. For example, for this reason alone, Catalonia's delay compared to Madrid is approximately one school year.

**Keywords:** immersion programmes, academic achievement, educational equity, educational policy, educational assessment, structural equation modelling.

## Resumen:

La cuestión lingüística en el ámbito escolar presenta en España una dimensión educativa de notable interés social, toda vez que un proceso inmersivo franco en las lenguas regionales podría estar afectando al rendimiento de los alumnos en función de su extracción lingüística y sociocultural. En este contexto, el presente trabajo tiene como objetivo principal analizar las relaciones entre nivel socioeconómico y cultural (ISEC), porcentaje de alumnos que hablan en la escuela un idioma diferente al de su hogar (idioma) y rendimiento escolar en PISA (rendimiento). En particular, se pretende medir el poder predictivo de la segunda variable con respecto a la tercera desde una perspectiva territorial. Se ha desarrollado una metodología basada en análisis de regresión lineal bivaria-

da y en modelos de ecuaciones estructurales aplicadas sobre datos de esas tres variables, disponibles en las bases de PISA 2015 y PISA 2022. La realización de estos análisis secundarios ha confirmado: a) la importancia del poder predictivo del ISEC sobre el rendimiento de forma consistente, tanto en 2015 como en 2022; b) la intensificación del vínculo entre ambas variables en dicho período de tiempo; y c) la aparición de una asociación moderada, pero significativa, de la variable idioma como predictor del rendimiento en el ámbito territorial. Se aprecia que el impacto de las diferencias territoriales en materia de ISEC sobre el rendimiento ha aumentado de forma notable entre 2015 y 2022, lo que resulta compatible con el incremento de otras brechas territoriales descritas en la literatura. Por otra parte, los modelos de una inmersión lingüística severa marcan diferencias significativas entre territorios en materia de rendimiento que se sustancian, a modo de ejemplo, en un retraso, solo por este motivo, de Cataluña respecto de Madrid de un año escolar aproximadamente.

**Palabras clave:** programas de inmersión, rendimiento académico, equidad educativa, política educativa, evaluación educativa, modelos de ecuaciones estructurales.

## 1. Introduction

The question of language in the school setting (understood principally in the case of Spain as the problem associated with the coexistence in schools of the official state language and, where applicable, of the co-official language of the particular autonomous community) has become more relevant as

models based on families' freedom of choice (or, failing that, an accommodating coexistence between the two languages) have gradually been replaced by models of strong immersion in the co-official language. This process of immersion might be affecting students' performance depending on their linguistic or sociocultural background.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) some time ago established a doctrine (UNESCO, 1953) of favouring use of the mother tongue (the first language that the child learns) in early childhood and primary education. Later on, in its Global Education Monitoring Report (UNESCO, 2016), it warned that “international and regional learning assessments confirm that when home and school languages differ there is an adverse impact on test scores” (p. 24). A more recent document (UNESCO-IIEP, 2021), which provides a significant review of what is known in this regard (Benson, 2004; Bühmann & Trudell, 2007; Pinnock, 2009a, 2009b), states that

most research now concludes that learning achievement is enhanced when children are taught in their mother tongue for at least the first six years of primary school before the second language, the main language of instruction, is introduced. (p. 1)

Moreover, international research has argued that education based on the mother tongue is of particular benefit to disadvantaged sectors, including in this group those from rural settings (Ball, 2011; Hovens, 2002).

Regarding developing countries, Reily (2019) mentions a long list of experimental studies with control groups based on standardised tests, which show the benefits of using the mother tongue in school in a notable range of countries. The institutional documents cited above adopt a position that summarises the contributions of previous studies and reviews aligned with

similar results. These results are consistently repeated in the setting of the developed countries.

For example, in the United States, Thomas and Collier (1997, 2002) carried out a large-scale longitudinal study on educational provision for minority-language students in public schools in the USA and their effect on academic performance, based on the records of around 200 000 students. They found that the strongest predictor of performance in the second language is the amount of formal education in the mother tongue.

In the European context, the most notable position is that of the Council of Europe, which, despite its clear support for multilingualism, accepts and develops the position established by UNESCO with regards to the use of the mother tongue in school. This is set out in Document 10837, of 7 February 2006, of the Parliamentary Assembly, which includes the report of the Culture, Science and Education Commission (Parliamentary Assembly, 2006) and which in point B.12 of its explanatory memorandum takes the following position:

Obviously, in European societies, everyday use of the official language is the main precondition for the integration of children whose main language is different from the official one of the country or region. However, a great deal of research is agreed on one point: immediate schooling of such children in a language they do not know well or at all (here, the term is teaching “through submersion”) seriously jeopardises their chances of academic success. Conversely, bilingual education based

on the mother tongue is the basis for long-term success. (p. 2)

It is important to emphasise that the results described above do not in any way invalidate the desirable objective (Bialystok *et al.*, 2012) of multilingualism in school, but that they simply question the use of some procedures or strategies for achieving it. In this sense, Cummins (2001), on the basis of the research, underlines the following:

Bilingualism has positive effects on children's linguistic and educational development. When children continue to develop their abilities in two or more languages throughout their primary school years, they gain a deeper understanding of language and how to use it effectively. ... The research suggests that bilingual children may also develop more flexibility in their thinking as a result of processing information through two different languages. (p. 17)

Nonetheless, Cummins (2001) aligns himself with the positions mentioned above relating to the use of the mother tongue in school when he says that “mother tongue promotion in the school helps develop not only the mother tongue but also children's abilities in the majority school language” (p. 18). This is what he calls “interdependence theory”.

In the specific case of Spain, studies on linguistic teaching models' effects on students are scarce (Clots-Figueras & Masella, 2013; Calero & Choi, 2019; Carabaña & Fernández-Enguita, 2019), particularly relating to their impact on school per-

formance. The work by Calero and Choi (2019) mentioned above which centres on academic performance and is solely limited to the case of Catalonia is notable in this regard. The authors perform secondary analyses based on the PISA 2015 database using multilevel linear models, and they conclude that the immersion policy has a negative effect on academic performance which they summarise as follows: “The results we have set out clearly point to the existence of a problem of equity, that generates ‘losers’ from the policy of linguistic immersion in Catalonia” (p. 16, own translation).

Moreover, the work of Carabaña and Fernández-Enguita (2019) centres on the relationship between regional languages in Spain and students' sense of belonging in the educational centre. To do this they use secondary analyses of the PISA 2015 database and conclude that this feeling (which is relevant both for students' well-being and for their performance) is notably lower in the bilingual autonomous communities, as “all of the students who study in regional languages depress the level of identification with the centre, regardless of the language of their homes” (p. 1, own translation). The authors propose as an explanatory hypothesis a “school-society diglossia”: the sense of belonging to the school suffers when the language of instruction is the one that is less preferred or of less social use.

Although not directly related to academic results, the work of Clots-Figueras and Masella (2013) on language, education, and identity, based on data from surveys

carried out in Catalonia, shows that respondents who have been exposed for longer to teaching in Catalan have stronger Catalan feelings. The effect also seems to be present among individuals whose parents are not of Catalan origin; furthermore, according to the authors, the reform also affects political preferences and attitudes towards the organisation of the state.

The recent publication of the PISA 2022 data, with a notable fall in scores for Catalonia in the three main areas of the assessment (OECD, 2023), has again put the strictly educational dimension of the linguistic question on the agenda in the social and political spheres (López-Rupérez & García-García, 2023; Rodríguez, 2023; Rodríguez-Sanmartín, 2023).

However, this is a question that concerns Spain as a whole, as well as Catalonia, for different reasons. Firstly, because we are facing situations that affect fundamental rights that the state is responsible for guaranteeing. And also, because seven autonomous communities (at present, the Balearic Islands, Catalonia, the Valencian Community, Galicia, Navarre, and the Basque Country) have models where co-official languages coexist with Spanish, which is the official language of the whole of Spain. In some cases, these linguistic models have approached broad immersion. It is, therefore, advisable to find evidence relating to the case of Spain that makes it possible to reflect on the magnitude of the effects and their consequences on an empirical basis.

For these reasons, in the present work, this question is approached from a territorial perspective, adopting a focus that combines the socio-economic and academic performance aspects with the linguistic ones, using the PISA data to do so. The main objectives of this research are:

1. To provide a comparative overview, by means of a descriptive analysis, of the 2015 and 2022 PISA results, with a detailed presentation of the scores for: performance; Economic, Social, and Cultural Status (ESCS); and language in different autonomous regions and cities of Spain.
2. To construct and evaluate a structural equation model (SEM) of the influence of the ESCS index on academic performance in science, mathematics, and reading on the sample of data from PISA 2015.
3. To analyse the impact of the language spoken at home on academic performance in PISA 2022 by means of a bivariate linear regression analysis.
4. To build and evaluate an SEM (structural equation model) of the relationship between ESCS and the percentage of students who speak a different language to the one that is dominant in schools (language) and its influence on academic performance in science, mathematics, and reading in the sample of data from PISA 2022.



## 2. Method

### 2.1. Samples

The present study uses two data sets. On the one hand, it uses the national PISA sample (OECD, 2016) corresponding to the 2015 edition. This sample comprises 6736 students (39 066 when including the expanded samples from the autonomous communities) and is representative of the Spanish population of 15-year-old stu-

dents, mainly enrolled in the fourth year of obligatory secondary education. It also uses the sample from the 2022 edition of PISA (OECD 2023) in a similar way. This comprises 30800 students as it considers all of the available samples from the autonomous communities and cities.

Table 1 presents detailed information about the sizes of the samples used according to the different regions.

TABLE 1. Size and representativeness of the sub-national samples in the 2015 and 2022 Editions of PISA.

	2015		2022	
	<i>n</i>	%	<i>n</i>	%
Spain	6736	17.2	-	-
Andalusia	1813	4.6	1707	5.5
Aragon	1798	4.6	1420	4.6
Asturias	1790	4.6	1724	5.6
Balearic Islands	1797	4.6	1576	5.1
Canary Islands	1842	4.7	1686	5.5
Cantabria	1924	4.9	1677	5.4
Castile and Leon	1858	4.8	1733	5.6
Castile-La Mancha	1889	4.8	1671	5.4
Catalonia	1769	4.5	1597	5.2
Valencian Community	1625	4.2	1625	5.5
Extremadura	1809	4.6	1685	5.9
Galicia	1865	4.8	1824	4.5
Madrid	1808	4.6	2138	6.9
Murcia	1796	4.6	1669	5.4
Navarre	1874	4.8	1826	5.9
Basque Country	3612	9.2	3244	10.5
La Rioja	1461	3.7	1394	5.3
Ceuta	-	-	345	1.1
Melilla	-	-	259	0.8
Total	39 066	100	30 800	100.0

Source: prepared by the authors based on the PISA 2015 and PISA 2022 databases.

## 2.2. Variables

The variables of interest have been organised into three basic categories, treating them as latent variables. In a context like the present one where SEM is used, it is important to note that latent variables are underlying constructs that are not directly observable but instead are inferred through observed variables. The three categories of latent variables used in the study are:

Performance latent variable (performance), disaggregated by areas:

- Mean performance in mathematics: this includes the gross values for the mathematics subject evaluated in 2015 (Per\_Mat\_15) and in 2022 (Per\_Mat\_22).
- Mean performance in language: this includes the gross values for the language subject evaluated in 2015 (Per\_Rea\_15) and in 2022 (Per\_Rea\_22).
- Mean performance in science: this includes the gross values for the science subjects evaluated in 2015 (Per\_Sci\_15) and in 2022 (Per\_Sci\_22).

Economic, cultural, and social status (ESCS) index latent variable:

- ESCS of the values reported in PISA 2015 (ESCS\_15).
- ESCS of the values reported in PISA 2022 (ESCS\_22).

Percentage of students who speak a different language than that of the school at home (language) latent variable:

- Percentage of students who speak a different language at home and at school, according to the figures from the 2022 PISA report (Language\_22).

## 2.3. Analytical procedures

In line with the objectives of the research, we have performed an analysis with three basic focusses. As a first approach, a comparative analysis of relevant data with an essentially descriptive orientation was performed. For this comparison, the mean values of the variables described above corresponding to the years 2015 and 2022 were compiled. To complement the above, a bivariate linear regression analysis was performed to assess the predictive value of the language variable with regards to the performance variable for 2022. Finally, various SEM structural methods were used (Bryman & Cramer, 1994; Pérez *et al.*; 2013) to validate a causal model that jointly encompasses the performance and ESCS variables in 2015, and another identical model for 2022, in this case adding the language latent variable to the analyses.

The structural equations model used derives from a technique that combines factor analysis with linear regression, going beyond simple measurement of the direct impact of the independent variables on the dependent one by considering the interaction and possible indirect influence between them. Although it does not demonstrate strict causality, it facilitates reasoning about causal hypotheses that are supported by empirical evidence (Batista *et al.*, 2000). The assumptions in

this research of the existence of at least one weak causal order and of the causal closure are accepted (Kim & Kohout, 1988). This type of approach is a common practice in studies of a similar nature (Bulut *et al.*; 2012; Calero & Choi, 2019; López-Rupérez *et al.*; 2021; Salas *et al.*; 2017; Swe *et al.*; 2020).

Version 2.3 of the *Jamovi* statistical analysis software was used (The jamovi project, 2022) to perform these analyses. This is based on the R lavaan software package (Rosseel, 2012) for structural equation modelling (Gallucci & Jentschke, 2021) and for graphic representation of path diagrams (Epskamp *et al.*; 2019).

### 3. Results

#### 3.1. A comparative overview of the PISA 2015 and 2022 results

Table 2 shows the performance results from the 2015 and 2022 editions of PISA for by areas and mean, ESCS, and language. These data provide a comparative overview of academic performance and associated factors in different regions, making it possible to approach a comprehension of the educational and social dynamics over a sufficiently long time period. Furthermore, this period deliberately disregards the anomaly for Spain of the 2018 edition in the implementation of the tests (and consequently in their results), as PISA has done in the country reports (OECD, 2023).

TABLE 2. Performance, ESCS, and language scores of the samples in the national and sub-national OECD levels corresponding to the 2015 and 2022 PISA editions.

Zone	Per_Mat		Per_Rea		Per_Sci		Per_Mean		ESCS		Language
	2015	2022	2015	2022	2015	2022	2015	2022	2015	2022	2022
OECD	490	472	493	476	493	485	493	478	0	0	11.18
Spain	486	473	496	474	493	485	491	477	-.51	-.03	19.40
Andalusia	466	457	479	461	473	473	473	464	-.87	-.18	6.31
Aragon	500	487	506	488	508	499	505	491	-.39	.00	10.66
Asturias	492	495	498	497	501	503	497	498	-.42	.09	4.48
Balearic Islands	476	471	485	472	485	480	482	474	-.65	-.05	51.80
Basque Country	492	482	491	466	483	480	489	476	-.25	.18	15.30
Canary Islands	452	447	483	463	475	473	470	461	-.80	-.21	5.41
Cantabria	495	495	501	494	496	504	497	498	-.43	.03	3.82

## The impact of linguistic immersion and economic, social and cultural status on academic performance

Castile and Leon	506	499	522	498	519	506	516	501	-.44	.08	3.45
Castile-La Mancha	486	464	499	468	497	475	494	469	-.66	-.16	6.52
Catalonia	500	469	500	462	504	477	501	470	-.35	-.02	55.54
Valencian Community	485	473	499	482	494	483	493	479	-.53	-.10	24.56
Extremadura	473	469	475	468	474	479	474	472	-.79	-.14	2.32
Galicia	494	486	509	485	512	506	505	493	-.52	.07	36.19
La Rioja	505	493	491	487	498	500	498	493	-.46	-.06	8.17
Madrid	503	494	520	496	516	502	513	497	-.01	.25	5.82
Murcia	470	463	486	468	484	482	480	471	-.82	-.24	7.40
Navarre	518	492	514	478	512	489	515	486	-.32	.02	27.75
Ceuta	-	395	-	404	-	410	-	403	-	-.47	26.41
Melilla	-	404	-	405	-	414	-	408	-	-.79	15.92

Source: prepared by the authors based on OECD 2016 and 2023.

A systematic analysis of the data from Table 2 reveals the following:

- Regarding the performance variables, with the exception of Asturias, Cantabria, and La Rioja, which improve in at least one of the areas, all of the others worsen in all three. Accordingly, when the mean scores are taken into account, all of them apart from the first two worsen. Catalonia with 31 PISA points and Castile-La Mancha with 25 are the ones that display the largest declines.
- In the case of ESCS, all of the autonomous communities improve in relation to the OECD mean in the period between 2015 and 2022, which results

in Spain improving by .48 points, positioning itself above the updated mean. The improvements in Andalusia and the Balearic Islands stand out, with increases of .69 and .60 ESCS points respectively. This phenomenon is analysed in the discussion section.

- Regarding the percentage of students who speak a different language at home to that of the school, Catalonia and the Balearic Islands, with values of 55.54% and 51.80% respectively, are the autonomous communities with the highest values for this variable. At the other extreme, Extremadura and Castile and Leon, with 2.32% and 3.45% respectively, have the lowest values.

### 3.2. Structural equations model corresponding to PISA 2015

In a first approach, a model was defined that centred exclusively on the 2015 data, establishing the relationship between ESCS and performance in science, mathematics, and reading. Application of the SEM using the maximum likelihood method with non-linear optimisation evaluated a model that incorporates latent variables (ESCS and performance) and observed variables. This model reveals significant relations between ESCS\_15 and academic performance, supported by statistical tests and with an adequate fit.

Various indices were used to evaluate the quality of the model's fit, the results of which are shown in Table 3; these indices indicate that the proposed model displays a reasonable fit, significantly better than a reference model ( $\chi^2 = 92.48$ ,  $p < .001$ ). The indices, such as the standardised root mean squared residual (SRMR  $< 0.05$ ), the root mean square error of approximation (RMSEA  $< 0.05$ ), the adjusted goodness of fit index (AGFI  $> 0.95$ ), the comparative fit index (CFI  $> 0.90$ ), or the Tucker–Lewis index (TLI  $> 0.70$ ), support the solidity of the model, following the recommendations of Schermelleh-Engel *et al.* (2003).

TABLE 3. Indices of fit of the SEM PISA 2015 model.

	Values obtained
Value of $p$ in $\chi^2$	.008
$\chi^2/df$	4.87
GFI	1
AGFI	0.999
SRMR	0.047
RMSEA	0.009
CFI	0.91
TLI	0.731
Bentler–Bonett Normed Fit Index (NFI)	0.895
Parsimony Goodness of Fit Index (PGFI)	0.143
Parsimony Normed Fit Index (PNFI)	0.298
Akaike Criterion (AIC)	378.019

Table 4 shows the parameter estimations, along with the measurement model for the relationship between academic performance and ESCS in 2015. The first section shows the path coefficient between academic performance and ESCS, indicating that an increase of 1 ESCS unit is associated with an estimated increase of 48.3 units in academic performance, with a confidence interval of 95% between 28.1 and 68.4. This coefficient is significant ( $\beta = 0.748, p < .001$ ), suggesting a positive relationship between ESCS and academic performance.

Similarly, Table 4 shows the estimates for the model for measuring the latent variables. ESCS\_15, as the observed variable of

the ESCS latent variable, has a standardised coefficient of 1.000, which indicates a perfect relationship. Likewise, with regards to academic performance, it can be seen that the observed variables (Per\_Sci\_15, Per\_Mat\_15, Per\_Rea\_15) have standardised coefficients of 1.00, 0.99, and 0.90 respectively, signalling the relationships existing between these observed variables and the latent variable of performance.

These standardised coefficients, along with their confidence intervals and significance values, provide information about the quality and strength of the relationships modelled. In general, Table 4 empirically supports the validity of the proposed model.

TABLE 4. Estimates of parameters and measurement model for the relationship between academic performance, language, and ESCS in the PISA 2015 results.

Dependent variable	Predictor variable	Estimate	Standard error	CI 95% (low/high)		$\beta$	$z$	$p$
Performance	ESCS	48.3	10.3	28.1	68.4	0.748	4.70	<.001
	<b>Latent variable</b>							
ESCS	<b>Observed variable</b>							
	ESCS_15	1.00	0.00	1.00	1.00	1.000		
	Per_Sci_15	1.00	0.00	1.00	1.00	0.992		
Performance	Per_Mat_15	0.99	0.14	0.71	1.27	0.861	6.84	<.001
	Per_Rea_15	0.90	0.07	0.76	1.03	0.965	12.69	<.001

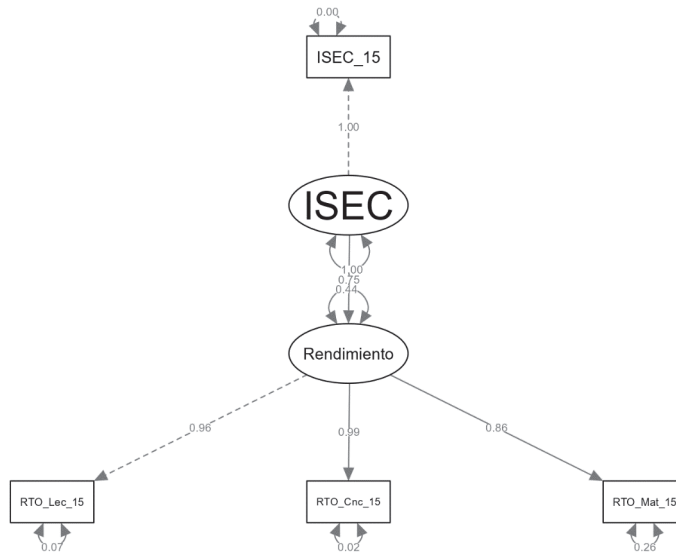
Figure 1 presents the path model illustrating the impact of ESCS on performance in PISA 2015. The SEM with data from PISA 2015 underlines the importance of

ESCS in predicting academic performance, supporting the hypothesis that higher ESCS is positively linked to better performance in science, mathematics, and reading.





FIGURE 1. Path model of the relationship between ESCS and academic performance in PISA 2015.



### 3.3. The predictive value of the language variable with regards to performance in PISA

A bivariate linear regression analysis was performed to analyse the impact of the language variable on academic performance in PISA 2022. Table 5 presents

key metrics that illustrate the connection between the two variables. The multiple correlation coefficient displays a weak relationship, while the low  $R^2$  values indicate that only a small proportion of the variability in performance is attributed to language.

TABLE 5. Regression statistics.

Multiple correlation coefficient	0.17
$R^2$ coefficient of determination	0.03
$R^2$ adjusted	-0.02
Standard error	26.14
Observations	21

Table 6 shows the global evaluation of the model in greater depth. The lack of statistical significance ( $F = 0.59$ ,  $p > .05$ ) and the modest sum of squares in the regression point to a generally weak direct

relationship between language and performance. Furthermore, the coefficients associated with the intercept and the language variable did not achieve sufficient significance.

TABLE 6. Analysis of variance and coefficients.

	Degrees of freedom	Sum of squares	Mean of the squares	F	p
Regression	1	404.42	404.42	0.59	0.45
Residuals	19	12984.62	683.40		
Total	20	13389.03			

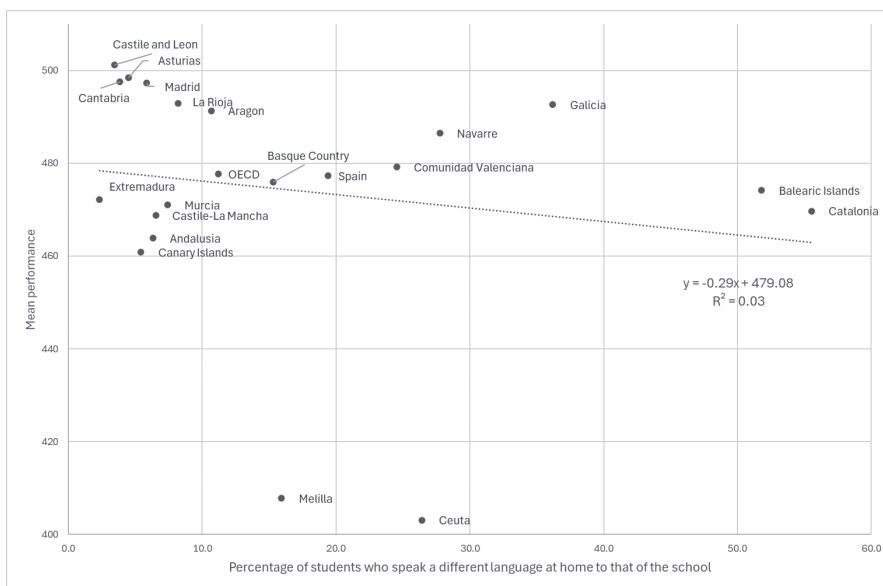
  

	Coefficient	Standard error	CI 95% (low/high)	t	p
Intercept	479.07	8.47	461.33 496.81	56.53	0.00
Variable X 1	-0.29	0.38	-1.08 0.50	-0.77	0.45

Although the Language variable does not show a statistically significant impact on the PISA 2022 performance scores, graphical analysis of this weak relation-

ship (see Figure 2) shows an inverse trend between the two variables: performance in PISA 2022 and use of language in samples at OECD, national, and sub-national levels.

FIGURE 2. Linear regression analysis of PISA 2022 performance vs language on the OECD, national, and sub-national samples.



### 3.4. Structural equation model corresponding to PISA 2022

Finally, the SEM focus applied to 2015 was replicated by means of a model that operates on the results of PISA 2022. This analysis again establishes a relationship between ESCS and performance in science, mathematics, and reading, but the results relating to the percentage of students who speak a language at home that is different to the one used at school (language) are also included. Although a non-significant relationship was established in the previous section, interpreted as absence of direct impact, it is important to note that an SEM has the capacity to model more complex relationships between variables, considering indirect effects and mediations.

The application of the SEM analysis using the maximum likelihood method

with non-linear optimisation has evaluated a model that includes latent variables (ESCS, language, and performance) and observed variables. This model finds significant relationships between ESCS\_22 and academic performance, as well as between the percentage of students who speak a different language to that of the schools (language), supported by statistical tests and an adequate fit.

In relation to the fit of the model, the indicators in Table 7 show that the proposed model displays a reasonable fit, considerably higher than that of a reference model ( $\chi^2 = 204.6$ ,  $p < .001$ ). Indices such as SRMR ( $< 0.05$ ), RMSEA ( $< 0.05$ ), AGFI ( $> 0.95$ ), CFI ( $> 0.90$ ), and TLI ( $> 0.85$ ) support the robustness and validity of the model.

TABLE 7. Fit indices of the SEM model in PISA 2022.

	Values obtained
Value of $p$ in $\chi^2$	.01
$\chi^2/df$	2.65
GFI	1
AGFI	0.999
SRMR	0.04
RMSEA	0.014
CFI	0.939
TLI	0.869
NFI	0.909
PGFI	0.259
PNFI	0.424
AIC	642.657

Table 8 shows the parameter estimates as well as the measurement model for the relationship between academic performance, language, and ESCS in 2022. In the section relating to academic performance, an increase of 1 ESCS unit is seen to be associated with an estimated increase of 166.78 units in academic performance, with a confidence interval of 95% between 98.78 and 234.78. This coefficient is significant ( $p < .001$ ), indicating a positive relationship

between ESCS and academic performance. Likewise, for the language predictor, it is found that an increase of 1 unit in the percentage of students who speak a language other than that of the school is related with an estimated reduction of 0.43 units in academic performance, with a confidence interval of 95% between -0.81 and -0.04. This coefficient is also significant ( $p < .05$ ), signalling a negative relationship between language and performance.

TABLE 8. Estimates of parameters and measurement model for the relationship between academic performance, language, and ESCS in the PISA 2022 results.

Dependent variable	Predictor variable	Estimate	Standard error	CI 95% (low/high)		$\beta$	$z$	$p$
Performance	ESCS	166.78	34.69	98.78	234.78	1.037	4.81	<.001
Performance	Language	-0.43	0.20	-0.81	-0.04	-0.245	-2.16	<.05

Latent variable	Observed variable	Estimate	Standard error	CI 95% (low/high)	$\beta$	$z$	$p$
ESCS	ESCS_non_immig_22	1.00	0.00	1.00	1.00	0.805	
	ESCS_immig_22	1.53	0.38	0.78	2.28	0.723	3.99 <.001
Language	%_languages_22	1.00	0.00	1.00	1.00	1.000	
	Per_Mat_22	1.00	0.00	1.00	1.00	0.977	
Performance	Per_Rea_22	0.92	0.05	0.82	1.02	0.992	18.18 <.001
	Per_Sci_22	0.94	0.05	0.85	1.05	0.995	19.09 <.001

With regards to the latent and observed variables, it should be noted that the relationships between ESCS and its components, as well as between language and the percentage of students who speak different languages, are significant and coherent with the model proposed. The observed variables of

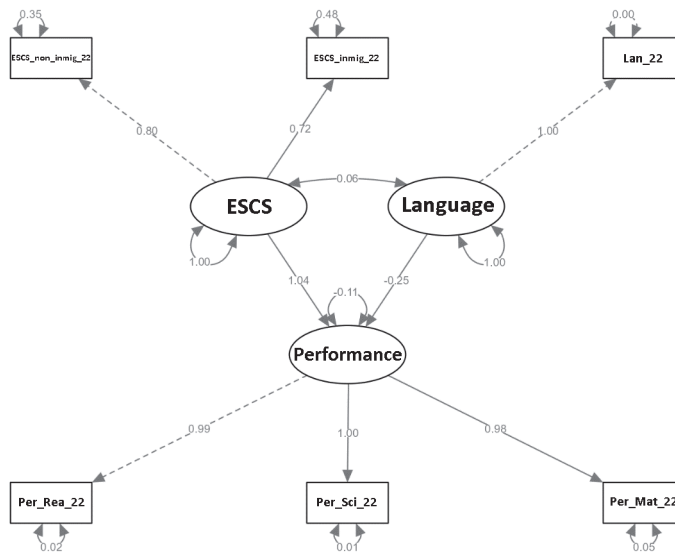
performance in the specific subjects also display positive and significant relationships with the latent performance variable. As a whole, these detailed estimates offer a deeper understanding of the interactions between the key variables in the context of the SEM model applied to the year 2022.



In summary, regarding the specific relationships from PISA 2022, a positive and strong association has been observed between ESCS and performance, as is a negative relationship of moderate but sig-

nificant intensity between language and academic performance. These conclusions are supported visually by the path model in Figure 3.

FIGURE 3. Path model of the relationship between ESCS and language with academic performance in PISA 2022



## 4. Discussion

As a preliminary question, we must underline here that the present work is aligned with re-evaluation of the path analysis [proposed by Judea Pearl (Pearl & Mackenzie, 2020)] as a procedure that goes beyond a mere statistical exercise, even though it contains one. It is based on an understanding of the problem by each scientist that is reflected in a causal diagram and is then submitted to the test of whether it matches reality, measured by the data and by quantitative analyses (pp. 99–100).

### 4.1. The influence of ESCS on performance

When analysing the evolution of ESCS between 2015 and 2022 and interpreting the advances experienced by all autonomous communities and by Spain as a whole, it is necessary to take into consideration the fact that ESCS is a normalised indicator (based on the set of the member states of the OECD) with a mean value equal to 0 and a standard deviation equal to 1 (OECD, 2023). In this time period, four countries with a lower level of development were added [Latvia in 2016,

Lithuania, and Colombia in 2018, and Costa Rica in 2021] something that would, at least partly, explain why Spain's ESCS value has improved relative to the mean.

On the other hand, and given the predictive value of the ESCS on performance, this would need to have resulted in a corresponding improvement in results with regards to the OECD mean, something that is not observed (see Table 2). This plausibly indicates that other factors with the opposite value have been at play in this period.

The results derived from the first model, despite being based on clusters of data grouped by autonomous communities, are coherent with the empirically established relationship in the successive editions of PISA that confirms the predictive value of ESCS in relation to performance. For example, the value of the correlation coefficient between these two variables established in PISA 2015 for the group of all participating countries was 0.129 (OECD, 2016). Our own linear regression analysis, limited to Spain in the area of science, gave a value of  $R = 0.38$  (López-Rupérez *et al.*, 2018, 2019a).

According to the final PISA 2022 report (OECD, 2023), 20 PISA points represent the “annual pace of learning” and so the figures above ultimately reflect estimated delays in the order of two years for each ESCS point (according to Table 4, one ESCS unit is associated with an estimated increase of 48.3 PISA units in academic performance). Even when the figure for the pace of the lag that the OECD provides is

the mean of the participating countries, it provides a basis to make an approximate estimate of from which observed differences regarding ESCS between autonomous communities are predicted (see Table 2). So, for example, the differences between Andalusia (-0.86) and the Community of Madrid (-0.01), in line with the results described above, would represent imbalances between the two regions equivalent to around two school years. This indicates a need for state-level policies regarding inter-regional equity.

#### 4.2. The results of using SEM (structural equation modelling)

The analysis of the predictive value of the percentage of students who speak a language at home other than the one they use at school with regards to the average scores in PISA 2022, has provided insufficient results. Consequently, if other intervening variables are not considered, the language variable on its own has limited predictive value for performance in PISA. Nonetheless, it is crucial to note that more detailed analyses, such as those using SEM models, have modified this first conclusion, enabling a more complete understanding of the underlying relationships.

Applying the SEM model to PISA 2022 not only replicates the relationships identified in 2015, but it also expands comprehension when including the language variable, thus enriching the analysis by exploring how linguistic dynamics can influence academic performance. This inclusion strengthens the empirical basis compared with the 2015 model, and also implicitly introduces the time variable by



considering the situation seven years later. According to the results mentioned above, the impact of the differences between autonomous communities, in terms of socio-economic and cultural level measured by the ESCS in 2022, has multiplied by 3.5 compared with 2015. In relative terms, this suggests a loss of inter-regional equity in Spain's education system.

This conclusion is compatible with the evolution in the territorial imbalances observed in other studies for variables with similar and different results. For example, the imbalance relating to the impact parameter (measured by the slope of the performance *vs.* ESCS regression line) increased from 11.3 to 14 points between 2015 and 2022 (López-Rupérez *et al.*, 2018, 2019b; López-Rupérez & García-García, 2023). Although it refers to different variables, a study on early school leaving by Soler (2022), carried out in the 2005–2021 period with two different analysis methodologies, consistently revealed “a lack of convergence between the different regions of Spain in the period analysed, above all in the case of males” (p. 153, own translation).

Despite this tendency, previous research and the OECD's own analyses (OECD, 2023; Sanz & Tena, 2023) have identified the partial effect of the covid-19 pandemic on the fall in scores in PISA 2022. It is also possible to surmise, on the basis of available evidence, that this effect has not only been felt in the mean scores but that it has done so on disadvantaged sectors in particular, which would explain the notable increase

in the impact of ESCS on performance described in this research. So, the study by Woessmann *et al.* (2020) concluded that students from socio-economically disadvantaged backgrounds were most affected by the pandemic. Chetty *et al.* (2020) have established a reduction of 60% in advances in academic performance of the students from the lowest quartile of income. And Bacher-Hick *et al.* (2021) provided evidence that could explain the previous results, namely that the highest growth in use of online educational resources occurs in urban areas with higher incomes.

In addition to the statistically significant relations between ESCS<sub>22</sub> and academic performance, the model reveals the connection between this and the percentage of students who speak a language other than that of the school at home (language). These associations support the idea that socio-economic context and linguistic variables both contribute significantly to student performance.

Regarding the language predictor within the corresponding SEM model, we found that an increase of 1 unit in the percentage of students who speak a language other than that of the schools is related to an estimated reduction of 0.426 units in academic performance, with a moderate but negative association between language and performance that is statistically significant. This means that the percentage of students who speak a different language at school to the one they use at home results in differences in academic performance measured by means of the PISA tests, thus contributing to generating gaps in

performance between autonomous communities. Although other factors will probably be at play in the relationship between these two variables, there is no doubt that in this order of ideas, full immersion linguistic models are associated with a significant difference between the regions in performance when the socio-economic and cultural level is considered. For example, the difference in these percentages between Catalonia and Madrid is approximately 50 percentage points (see Table 2, Language), which, according to the SEM model, corresponds to 21 PISA academic performance points and results in an academic delay for this reason alone (according to Table 2, the total difference is 27 PISA points) of something more than one academic year between these two autonomous communities. This highlights the advisability of reviewing models of linguistic immersion in school, taking into account their impact on academic performance.

## 5. Conclusions

Despite the sophistication of the SEM model used, the correlation identified between the variables does not necessarily imply direct causality. These limitations open the door to future research that tackles these methodological restrictions. Moreover, although the model considers key variables such as ESCS, language, and performance, other potentially influential factors have been omitted, such as the quality of teachers (Castro & Egido, 2024), the quality of school leadership (Gairín, 2024), and other empirically relevant factors (Hattie, 2023) that, along with the previous variables, might affect academic performance.

For this reason, a more complex model incorporating additional variables would be a fruitful direction for future research.

This necessary caution, along with the recommendation for future research above, do not prevent us from noting that the results of this research agree with others already consolidated in the specialist literature, some of which are cited above.

So, the models used for both the 2015 and 2022 editions of PISA suggest that an increase in ESCS is associated with better academic performance at the time of the evaluation, as a significant relationship was consistently found between a higher ESCS and better academic performance. This agrees with the repeated and positive association between students' socio-economic and cultural level and their academic performance that the different editions of PISA have repeatedly shown (OECD, 2016, 2023), as has another research relating to Spain (López-Rupérez *et al.*, 2019c).

When using a model that takes into account the ESCS variable as well as the language and performance variables, speaking a language at school other than the one used at home can, in strong immersion models, have a significant negative impact on academic results from PISA 2022. This is consistent with the extensive body of research that has confirmed that education based on the mother tongue improves academic results and significantly increases chances of educational success, as underlined by UNESCO (2021) and the Council of Europe (Parliamentary Assembly, 2006) among other institutions.

Ultimately, and beyond arguments that can be invoked alongside education, applying this principle of precaution [formulated in accordance with the classical *dictum*, *primun non nocere* ('first do no harm') and which is not only a deontological requirement in the field of health but also in education] points to the need to align educational policies and their results with the demands of equity and of the best interest of the child.

## Authors' contributions

**Francisco López-Rupérez:** Conceptualisation; Project administration; Supervision; Writing (original draft).

**Álvaro Moraleda-Ruano:** Conceptualisation; Data curation; Visualisation; Writing (original draft).

**Isabel García-García:** Data curation; Visualisation.

## References

- Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, 193, 104345. <https://doi.org/10.1016/j.jpubeco.2020.104345>
- Ball, J. (2011). *Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000121270>
- Benson, C. (2004). *The importance of mother tongue-based schooling for educational quality*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000146632>
- Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16 (4), 240-250.
- Bryman, A., & Cramer, D. (1994). *Quantitative data analysis for social scientists*. Routledge.
- Bühmann, D., & Trudell, B. (2007). *Mother tongue matters: Local language as a key to effective learning*. United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://unesdoc.unesco.org/ark:/48223/pf0000161121>
- Bulut, O., Delen, E., & Kaya, F. (2012). An SEM model based on PISA 2009 in Turkey: How does the use of technology and self-regulation activities predict reading scores? *Procedia - Social and Behavioral Sciences*, 64, 564-573. <https://doi.org/10.1016/j.sbspro.2012.11.066>
- Calero, J., & Choi, Á. (2019). *Efectos de la inmersión lingüística sobre el alumnado castellanoparlante en Cataluña [Effects of linguistic immersion on Spanish-speaking pupils in Catalonia]*. Fundación Europea Sociedad y Educación. <https://sociedadyededucacion.org/site/wp-content/uploads/SE-Inmersion-Cataluna.pdf>
- Carabaña, J., & Fernández-Enguita, M. (2019). Enseñanza en lenguas regionales y sentimiento de pertenencia a la escuela [Schooling in regional languages and sense of belonging at school]. *Journal of Supranational Policies of Education*, (9), 57-94. <http://doi.org/10.15366/jospoe2019.9.002>
- Castro, M., & Egido, I. (2024). *Qué sabemos sobre el profesorado. Políticas, evidencias y perspectivas de futuro [What we know about teachers. Policies, evidence and future perspectives]*. Narcea.
- Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). How did covid-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data. *NBER Working Paper*, (27431). <https://doi.org/10.3386/w27431>
- Clots-Figueras, I., & Masella, P. (2013). Education, language and identity. *The Economic Journal*, 123 (570), 332-357.

- Cummins, J. (2001). Bilingual children's mother tongue: Why is it important for education? *Sprogforum*, 7 (19), 15-20.
- Epskamp, S., Stuber, S., Nak, J., Veenman M., & Jorgensen T. D. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output* [Computer Software]. <https://CRAN.R-project.org/package=semPlot>
- Foguet, J. M. B., & Gallart, G. C. (2000). *Modelos de ecuaciones estructurales [Structural equation modelling]*. La Muralla.
- Gairín, J. (2024). *Dirección y liderazgo de los centros educativos. Naturaleza, desarrollo y práctica profesional [School management and leadership. Nature, development and professional practice]*. Narcea.
- Gallucci, M., & Jentschke, S. (2021). *SEMLj: jamovi SEM Analysis* [Computer Software]. <https://semlj.github.io/>
- Hattie, J. (2023). *Visible learning: The sequel. A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.
- Hovens, M. (2002). Bilingual education in West Africa: Does it work? *International Journal of Bilingual Education and Bilingualism*, 5 (5), 249-266.
- Kim, J. O., & Kohout, F. J. (1988). Special topics in general linear models. In N. H. Nie, D. H. Bent, & C. H. Hull, *Statistical package for the social sciences*. McGraw- Hill.
- López-Rupérez, F., García-García, I., & Expósito Casas, E. (2018). *PISA 2015 y las comunidades autónomas españolas. Diagnósticos y políticas de mejora [PISA 2015 and the Spanish autonomous communities. Diagnostics and improvement policies]*. Universidad Camilo José Cela. [https://www.ucjc.edu/pdfs/universidad/pisa2015\\_ccaa.pdf](https://www.ucjc.edu/pdfs/universidad/pisa2015_ccaa.pdf)
- López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019a). Rendimiento en ciencias, concepciones epistémicas y vocaciones STEM en las comunidades autónomas españolas. Evidencias desde PISA 2015, políticas y prácticas de mejora [Performance in science, epistemic conceptions, and STEM vocations in Spain's autonomous communities: evidence from PISA 2015, improvement policies, and practices]. *Revista Española de Pedagogía*, 272 (77), 5-27. <https://doi.org/10.22550/REP77-1-2019-09>
- López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019b). Educational effectiveness, efficiency, and equity in Spanish regions: What does PISA 2015 reveal? *ORBIS SCHOLAE*, 12 (2), 9-36. <https://doi.org/10.14712/23363177.2018.291>
- López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019c). Igualdad de oportunidades e inclusión educativa en España [Equal opportunities and educational inclusion in Spain]. *RELIEVE*, 25 (2), 1. <http://doi.org/10.7203/relieve.25.2.14351>
- López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España: análisis empíricos y recomendaciones políticas [Grade repetition and graduation in compulsory secondary education in Spain: Empirical analysis and policy recommendations]. *Revista de Educación*, 394, 325-354. <https://doi.org/10.4438/1988-592X-RE-2021-394-510>
- López-Rupérez, F., & García-García, I. (2023, 11 Diciembre). PISA 2022 y la cuestión territorial en España [PISA 2022 and the territorial question in Spain]. *Magisnet*. <https://www.magisnet.com/2023/12/pisa-2022-y-la-cuestion-territorial-en-espana/>
- OECD. (2016). *PISA 2015 results (volume I). Excellence and equity in education*. OECD Publishing. <https://doi.org/10.1787/9789264266490-en>
- OECD. (2023). *PISA 2022 results (volume I). The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Parliamentary Assembly. (2006). *The place of mother tongue in school education* (Doc. 10837). Committee on Culture, Science and Education. Council of Europe. <https://assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewHTML.asp?FileID=11142&lang=EN>
- Pearl, J., y Mackenzie, D. (2020). El libro del porqué. *La nueva ciencia de la causa y el efecto [The book of why. The new science of cause and effect]*. Pasado y Presente.
- Pérez, E., Medrano, L. A., & Sánchez-Rosas, J. (2013). El *path analysis*: conceptos básicos y ejemplos de aplicación [Path analysis: Basic concepts and examples]. *Revista Argentina de Ciencias del Comportamiento*, 5 (1), 52-66.
- Pinnock, H. (2009a). *Language and education. The missing link. How the language used in schools threatens the achievement of education for all*. CfBT Education Trust & Save the Children. <https://resourcecentre.savethechildren.net/pdf/1674.pdf>



- Pinnock, H. (2009b). *Steps towards learning: A guide to overcoming language barriers in children's education*. Save the Children. <https://www.savethechildren.org.uk/content/dam/global/reports/steps-towards-learning-lr.pdf>
- Reilly, C. (2019). *Language and development: Issues in mother tongue based multilingual education* [Inedit Manuscript]. [https://www.researchgate.net/publication/343768795\\_Language\\_and\\_development\\_Issues\\_in\\_mother-tongue\\_based\\_multilingual\\_education#fullTextFileContent](https://www.researchgate.net/publication/343768795_Language_and_development_Issues_in_mother-tongue_based_multilingual_education#fullTextFileContent)
- Rodríguez, P. (2023, 7 Diciembre). Catalunya se pregunta por qué se ha hundido más que el resto de España en PISA [Catalonia wonders why it has sunk further than the rest of Spain in PISA]. *Eldiario.es*. [https://www.eldiario.es/catalunya/catalunya-pregunta-hundido-resto-espana-pisa\\_1\\_10748162.html](https://www.eldiario.es/catalunya/catalunya-pregunta-hundido-resto-espana-pisa_1_10748162.html)
- Rodríguez-Sanmartín, O. (2023, 19 Diciembre). Los alumnos de Cataluña son los que muestran el mayor rechazo a la escuela de toda España [Pupils in Catalonia show the highest level of rejection of school in Spain]. *El Mundo*. <https://www.elmundo.es/espana/2023/12/19/6581fb-22fdddf570d8b45b5.html>
- Rosseel, &. (2012). Ijavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48 (2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Salas, A. A., Navarro, R., & Montero, E. (2017). Un modelo de ecuaciones estructurales para el estudio de factores que afectan la competencia lectora y la alfabetización matemática: una aproximación bayesiana con datos de PISA 2009 [A structural equation model for the study of factors affecting reading literacy and mathematical literacy: A Bayesian approach using PISA 2009 data]. *Estadística Española*, 59 (194), 167-192.
- Sanz, I., & Tena, J.D. (2023). The impact of the covid-19 pandemic on education learning. In J. Sanz, & I. Sanz (Eds.), *Addressing inequities in modern educational assessment* (pp. 15-36). Springer. [https://doi.org/10.1007/978-3-031-45802-6\\_2](https://doi.org/10.1007/978-3-031-45802-6_2)
- 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*, 8 (2), 23-74.
- Soler, A. (2022). Heterogeneidad regional en el abandono educativo temprano [Regional heterogeneity in early school dropout rates]. *Magisterio*, 24-25.
- Swe, K.M., Fraser, B. J., & Afari, E. (2020). Structural relationships between learning environments and students' non-cognitive outcomes: Secondary analysis of PISA data]. *Learning Environments Research*, 23 (3), 395-412. <https://doi.org/10.1007/s10984-020-09313-2>
- The jamovi project. (2022). *jamovi* (version 2.3) [Computer Software]. <https://www.jamovi.org>
- Thomas, W. P., & Collier, V. P. (1997). *School effectiveness for language minority students*. National Clearinghouse for Bilingual Education. [https://ncela.ed.gov/sites/default/files/legacy/files/rcd/BE020890/School\\_effectiveness\\_for\\_langu.pdf](https://ncela.ed.gov/sites/default/files/legacy/files/rcd/BE020890/School_effectiveness_for_langu.pdf)
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Center for Research on Education, Diversity and Excellence. <https://escholarship.org/uc/item/65j213pt>
- UNESCO. (1953). *The use of the vernacular languages in education*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO).
- UNESCO. (2016). *Si no entiendes, ¿cómo puedes aprender? Informe de seguimiento de la educación en el mundo* [If you don't understand, how can you learn? Global education monitoring report]. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). [https://unesdoc.unesco.org/ark:/48223/pf0000243713\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000243713_spa)
- UNESCO-IIEP. (2021, 23 de septiembre). *La lengua de enseñanza* [The language of teaching]. <https://bit.ly/3UsQE2C>
- Woessmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., & Zierow, L. (2020). Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen? [Education during the coronavirus crisis: How have schoolchildren spent the period of school closures and what educational measures are Germans in favour of?]. *ifo Schnelldienst*, 9, 25-39.



## Authors' biographies

**Francisco López-Rupérez.** He has a Doctorate in Physical Sciences from the Universidad Complutense de Madrid and a Secondary School Chair in Physics and Chemistry. He has been the director general of educational centres of the Ministry of Education, Culture, and Sports; secretary general for Education and Professional Training in this Ministry, and president of the School Council of the State. He has written a dozen books and over a hundred articles in Spanish and foreign journals. He is currently the Director of the Extraordinary Chair in Educational Policies of the Universidad Camilo José Cela and of the University Master's in Policies and Governance of Education Systems.

 <https://orcid.org/0000-0003-2613-9652>

**Álvaro Moraleda-Ruano.** He has a Doctorate in Pedagogy from the Universidad Complutense de Madrid, with special doctoral prize. His contribution to the Department of Theory and History of Education is reflected in various collaboration

grants, including an honorary grant. He has focussed his work in key areas such as emotional intelligence and statistical research methods in education. He is currently research coordinator in the Faculty of Education of the Universidad Camilo José Cela. He has an extensive record of scientific publications and has participated in various externally funded research groups and projects.

 <https://orcid.org/0000-0002-3638-8436>

**Isabel García-García.** She has a Licenciature degree in Mathematical Sciences from the Universidad Complutense de Madrid and holds a Secondary School Chair in Mathematics. She has been Head of Area of the School Council of the State and was responsible for the 2012 to 2016 editions of the *Informe sobre el estado del sistema educativo* [Report on the state of the education system]. She is currently a collaborator with the Chair in Educational Policies of the Universidad Camilo José Cela.

 <https://orcid.org/0000-0002-9884-6148>

