





Eficacia escolar y alto rendimiento lector del alumnado español en PISA 2018: un enfoque de machine learning

School effectiveness and high reading achievement of spanish students in PISA 2018: a machine learning approach

Delia Arroyo Resino ¹ 
Alexander Constante-Amores ^{2*} 
María Castro ¹ 
Enrique Navarro ¹ 

¹ Universidad Complutense de Madrid, Spain

² Universidad Camilo José Cela, Spain

* Autor de correspondencia. E-mail: israelalexander.constante@ucjc.edu

Cómo referenciar este artículo/ How to reference this article:

Arroyo Resino, D., Constante-Amores, A., Castro, M., & Navarro, E. (2024). Eficacia escolar y alto rendimiento lector del alumnado español en PISA 2018: un enfoque de machine learning [School effectiveness and high reading achievement of Spanish students in PISA 2018: a machine learning approach]. *Educación XX1*, 27(2), 223-251. <https://doi.org/10.5944/educxx1.38634>

Fecha de recepción: 14/10/2023
Fecha de aceptación: 01/02/2024
Publicado online: 28/06/2024

RESUMEN

En las últimas décadas, el análisis de la eficacia escolar ha adquirido una creciente importancia en el ámbito educativo. La presente investigación se centra en estudiar los factores de eficacia escolar asociados al alto rendimiento en la comprensión lectora. La muestra se encuentra conformada por los estudiantes españoles que participaron en PISA 2018. La variable criterio es el alto rendimiento en comprensión lectora y se ha contado con un total de 159 predictores relacionados con la eficacia escolar. Los datos se han analizado con el

algoritmo de Random Forest y regresión logística binaria multinivel. Entre los principales resultados se destaca que las variables más importantes son las variables de proceso: placer por la lectura y metacognición: evaluar la credibilidad. Además, se demuestra la importancia relativa que tienen los factores de contexto o entrada y proceso explicando un 41% y 38%, respectivamente, de la varianza de la variable criterio. El modelo final (formado por ambos grupos de factores) explica aproximadamente el 54% del éxito lector. En este modelo, el predictor que tiene un mayor efecto es la metacognición: evaluar la credibilidad, referido a la capacidad del sujeto para evaluar la calidad y credibilidad de un texto (por ejemplo, si la información es válida, precisa e imparcial), siendo su efecto aproximadamente el doble que el de las variables de contexto o entrada. Entre las principales conclusiones se destaca la posibilidad de aumentar el escaso número de estudiantes de alto rendimiento en el contexto español mediante el desarrollo de políticas educativas que fomenten el placer por la lectura y la capacidad metacognitiva.

Palabras clave: PISA, alto rendimiento, machine learning, eficacia escolar

ABSTRACT

In the last few decades, the analysis of school effectiveness has gained increasing importance in the field of education. This current research focuses on studying the factors of school effectiveness associated with high performance in reading comprehension. The sample is comprised of Spanish students who participated in PISA 2018. The dependent variable is high performance in reading comprehension, and a total of 159 predictors related to school effectiveness have been considered. The data were analyzed using the Random Forest algorithm and binary multilevel logistic regression. Among the key findings, it is highlighted that the most important variables are process variables: enjoyment of reading and metacognition: evaluating credibility. Furthermore, the relative importance of context or input factors and process factors explains 41% and 38%, respectively, of the variance of the criterion variable. The final model (comprising both groups of factors) explains approximately 54% of reading success. In this model, the predictor that has the most significant effect is metacognition: evaluating credibility, which refers to the subject's ability to assess the quality and credibility of a text (for example, whether the information is valid, accurate, and impartial), with its effect being roughly double that of context or input variables. Among the main conclusions, it is highlighted that it is possible to increase the scarce number of high-performing students in the Spanish context through the development of educational policies that promote a love for reading and metacognitive capacity.

Keywords: PISA, high achievement, machine learning, school effectiveness

INTRODUCCIÓN

La eficacia escolar es considerada un clásico dentro de la investigación educativa (Creemers et al., 2022), en este sentido, su estudio es una línea que ocupa un espacio propio en la literatura científica y lejos de ser una tradición agotada sigue siendo un ámbito fructífero y necesario de reflexión y de investigación (Scheerens & Creemers, 2022).

Su origen se remonta al Informe Coleman (1966) centrado en el estudio de la desigualdad en los logros académicos, donde se pone de manifiesto la importancia del contexto socioeconómico en los resultados educativos, por encima de las variables escolares (López-González et al., 2021). Según de la Orden et al. (1997) este fenómeno tiene que venir necesariamente expresado en indicadores que reflejen la relación entre los logros o resultados del sistema y las metas y objetivos planteados para el mismo. Por ello, la eficacia escolar no abarca solo el nivel de rendimiento obtenido por los estudiantes, las clases y los centros educativos individuales (calidad), sino también la distribución equitativa de los resultados de aprendizaje entre los estudiantes con diversas características de origen (equidad) (Kyriakides et al., 2019), por lo que tal y como indica Hu et al. (2021) es fundamental estudiar las peculiaridades de este grupo extremo, referido a los estudiantes con alto rendimiento, frente al resto. La sociedad, incluidos los organismos nacionales/ estatales y las escuelas/profesores, deberían asumir la responsabilidad de ofrecer oportunidades de aprendizaje y orientaciones adicionales a los grupos de estudiantes más desfavorecidos, de tal manera que se consiga una educación equitativa y de calidad (Creemers et al., 2022). De esta manera, los modelos de eficacia escolar se centran en el estudio de los procesos implementados por los centros educativos que contribuyen significativamente al rendimiento académico del alumnado, pues son los criterios básicos para juzgar la eficacia educativa. Así, la identificación de los factores educativos asociados al rendimiento estudiantil es un aspecto clave en la investigación educativa sobre la eficacia escolar (Creemers et al., 2022; Murillo, 2007; Scheerens et al., 2013).

Los resultados de las evaluaciones educativas internacionales, como el Programa Internacional de Evaluación de los Alumnos (PISA) que evalúan el rendimiento competencial de estudiantes de diferentes entornos educativos, sirven como indicadores estatales de eficacia (Kyriakides et al., 2019). Los cuestionarios aplicados a estudiantes, familias, profesorado y directores permiten avanzar en la investigación sobre la eficacia escolar, analizando los factores escolares que más afectan de manera extensa, precisa y rigurosa; lo que puede conducir a la mejora de los procesos y políticas educativas en las diferentes regiones evaluadas (Murillo, 2007), tema que aún no se ha explorado completamente con esta base de datos (Kyriakides et al., 2019).

Hasta el momento, la eficacia escolar se ha estudiado en PISA desde una doble perspectiva. En primer lugar, investigaciones que analizan este fenómeno desde la alta y baja eficacia escolar (Gamazo et al., 2018; Martínez-Abad et al., 2020). En segundo lugar, quienes investigan los factores más importantes que influyen en el rendimiento en ciencias de los estudiantes de alto y bajo rendimiento, como, por ejemplo, el estudio de Hu et al. (2021) quienes utilizan por primera vez en una evaluación internacional el modelo dinámico de eficacia escolar de Creemers y Kyriakides (2008). Sin embargo, actualmente no se han analizado los factores de eficacia escolar asociados al alumnado de alto rendimiento frente al resto de estudiantes, algo de gran interés para poder establecer estrategias que permitan mejorar la calidad de los sistemas educativos al aumentar los porcentajes de este tipo de estudiantes.

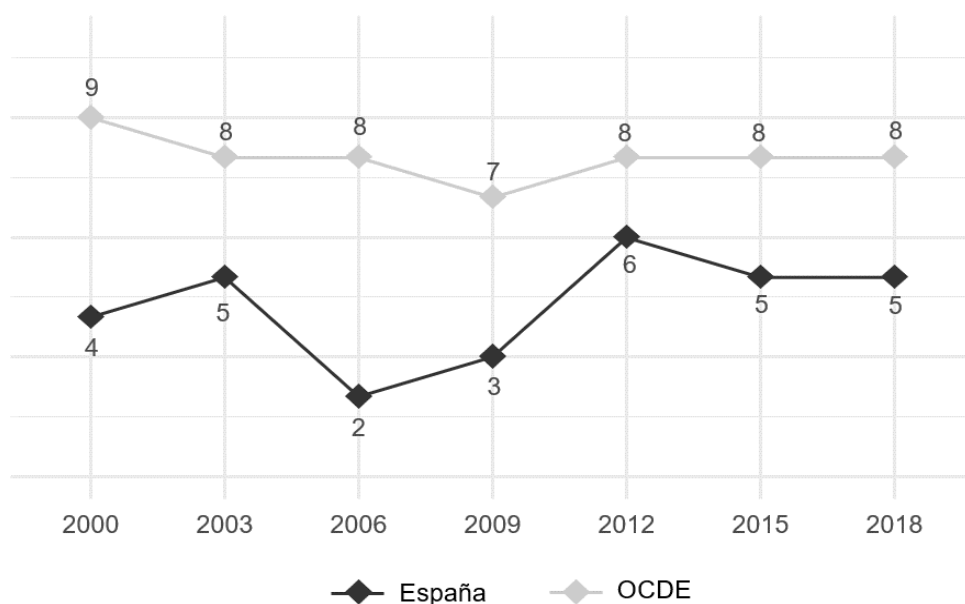
De las tres competencias troncales que evalúa de manera trienal PISA, la comprensión lectora es una herramienta esencial dentro del campo educativo, pues permite a los estudiantes acceder y comprender la realidad construyendo aprendizajes significativos. Además, esta es una habilidad fundamental, ya que continuamente nos encontramos produciendo y comprendiendo textos, se trata de una práctica cotidiana en el mundo actual y es la base de los aprendizajes autónomos en la sociedad del conocimiento (García et al., 2018; Molina, 2020).

En este sentido, parece existir una relación positiva entre la competencia lectora y los aprendizajes. García et al. (2018) encontraron una correlación positiva y significativa entre el nivel de comprensión de lectura de los estudiantes y el rendimiento en cuatro campos (Lengua Castellana, Matemáticas, Ciencias Sociales y Ciencias Naturales), algo que no es de extrañar, ya que las habilidades lectoras son una herramienta fundamental para la construcción de significados y conocimientos (Gómez et al., 2014). Por lo tanto, los estudiantes con un bajo nivel de comprensión lectora suelen tener un rendimiento bajo, ya que esta competencia afecta directamente a los aprendizajes (Viramontes et al., 2019).

La comprensión lectora se encuentra organizada en PISA 2018 en seis niveles, considerándose a los estudiantes de «alto rendimiento» aquellos que obtienen un nivel 5 y 6. Estos se caracterizan por ser lectores que pueden localizar, organizar e inferir información, realizar reflexiones críticas (nivel 5), hacer comparaciones, poseen una comprensión detallada de los textos, pueden integrar información de diferentes textos e incluso generar categorías abstractas (nivel 6) (OCDE, 2019). En España el porcentaje de estudiantes que se encontraba en estos niveles de «alto rendimiento» en el año 2018 era del 5%, mostrándose siempre por debajo del promedio de los países de la OCDE (8% en 2018), tal y como se muestra en la siguiente figura.

Figura 1

Evolución del porcentaje de estudiantes de alto rendimiento en lectura en PISA 2018



Fuente. Elaboración propia a partir de la base de datos de las diferentes ediciones de PISA.

Por lo tanto, uno de los principales problemas del sistema educativo español es que hay muy pocos estudiantes que se encuentran en los niveles altos de dicha competencia. Esto es realmente preocupante porque afecta directamente a la eficacia y a la equidad social, por lo que los sistemas educativos deberían poder incrementar el porcentaje de alumnos que se encuentran en estos niveles superiores de tal manera que se consiga el máximo rendimiento de los estudiantes (Gaviria, 2004).

Factores de eficacia asociados al rendimiento lector

El modelo de Contexto, Entrada, Proceso y Producto (CEPP) de Scheerens (1991) sintetiza la idea clásica de eficacia escolar. Existiendo una amplia literatura (Gamazo et al., 2018; Martínez-Abad, et al., 2020; Murillo, 2007) que organiza los factores de eficacia escolar que más comúnmente se asocian al rendimiento estudiantil en factores de contexto o entrada (características de los estudiantes, características de los docentes, infraestructuras, etc.), factores de proceso (actuación del profesorado en el aula y las tutorías, metodología didáctica, etc.) y factores de producto

(rendimiento académico). A su vez, estos factores pueden dividirse en dos niveles: estudiantes y centros educativos.

Factores de contexto o entrada

Entre los factores contextuales del alumnado que más se vinculan con el alto rendimiento en comprensión lectora, se encuentra un entorno socioeconómico y cultural favorecido (Campos y Arantegui, 2022). La investigación de Franco et al. (2016) refuerza la importancia que tiene en la comprensión lectora el número de libros en el hogar, los textos utilizados para realizar tareas escolares y la supervisión y el apoyo de los padres en la realización de estas.

También son variables importantes el sexo del estudiante y la condición de inmigrante. Con respecto a la primera, la investigación realizada por Frutos y Santaren (2020) demuestra una mejor comprensión lectora en alumnas que en alumnos, algo observable en todas las ediciones de PISA y PIRLS, aunque en PISA, la brecha de género se habría reducido en las últimas ediciones. En relación con la segunda, esta variable tiene una gran repercusión en el rendimiento lector. Concretamente los estudiantes autóctonos obtienen una mayor puntuación en lectura que los estudiantes inmigrantes de primera generación (inmigrantes que nacieron en otros países) (Cordero et al., 2013).

Referido a los factores propios del centro educativo que más se vincula con el alto rendimiento en comprensión lectora, dentro de las variables de contexto, se destaca la importancia de la titularidad de centro. Asensio-Muñoz et al. (2018) señalan que esta es la variable del centro educativo que tiene una mayor relación con el rendimiento en comprensión lectora y que los estudiantes que proceden de hogares socioeconómicamente desfavorecidos obtienen un menor rendimiento en esta competencia. Otra variable también importante es el hecho de haber cambiado dos o más veces de escuela, estos estudiantes tienen un menor rendimiento en lectura en comparación con aquellos que no han cambiado de escuela o lo han hecho una vez (Gamazo et al., 2018).

Factores de proceso

Entre los factores de proceso, destaca la importancia de las expectativas académicas del estudiante (Garrido et al., 2020 y Hu et al., 2021). El alumnado que espera completar estudios universitarios tiene una mayor probabilidad de obtener un alto rendimiento en la competencia lectora. La investigación de Franco et al. (2016), así como la de Herrera et al. (2017) indican la relevancia de un autoconcepto lector positivo, ya que repercute favorablemente sobre el rendimiento. Otra

variable que está relacionada con el éxito académico es la metacognición (Wu et al., 2020). En este sentido, un mayor conocimiento metacognitivo por parte del alumnado produce un mayor rendimiento en la competencia lectora en PISA (Artelt & Schneider, 2015). Específicamente, Qi (2021) señala que de las diversas estrategias metacognitivas, las relacionadas con el proceso de resumir textos son las más importantes. Otro factor de gran relevancia es el disfrute hacia la lectura, ya que los estudiantes que dedican más tiempo a la misma muestran un mayor rendimiento en dicha competencia (Akande y Oyedapo, 2018; Molina, 2020).

Respecto a los factores de proceso referidos al centro, una de las variables más importantes es el clima escolar (Linnakyla et al., 2004). También se destaca el uso de las TIC, jugando un papel fundamental en el alto dominio de la comprensión lectora, ya que actualmente nos encontramos sumergidos en la era de la tecnología y los estudiantes utilizan estas herramientas para leer, por lo tanto, se debe aprovechar esta situación para el fomento de la lectura (Rivera, 2013). Además, Avendaño y Martínez (2013) resaltan la importancia de las TIC para generar nuevos escenarios donde los estudiantes se aproximen de una forma innovadora a los textos e interactúen con ellos en el marco de la era digital, algo que ayuda en el desarrollo de la competencia lectora. Finalmente, otra variable que también destaca es la retroalimentación recibida por el profesorado (Hu et al., 2021), siendo beneficioso un feedback de calidad, equitativo en el tiempo.

A partir de lo comentado, el objetivo general de la presente investigación es estudiar los factores (contexto o entrada y proceso) de eficacia escolar, asociados al alto rendimiento en comprensión lectora en el marco de PISA 2018, utilizando la muestra española. Para su consecución, se formulan los siguientes objetivos específicos:

1. Identificar los factores de eficacia escolar que tienen una mayor influencia con el alto rendimiento lector.
2. Determinar la contribución relativa de los factores que más inciden en la probabilidad de que un estudiante obtenga un alto desempeño en la competencia lectora teniendo en cuenta la estructura jerárquica de los datos (nivel 1: estudiante y nivel 2: centro escolar).

MÉTODO

Se llevó a cabo un análisis secundario de los datos de la evaluación internacional de PISA 2018. Se trata de una investigación cuantitativa con un diseño no experimental basado en datos transversales y enmarcado dentro de los estudios ex post facto. Hay que destacar que en la presente investigación se siguió una doble estrategia metodológica. En primer lugar, se llevó a cabo un análisis descriptivo y exploratorio mediante el algoritmo de machine learning Random Forest. En segundo

lugar, se realizó un modelo predictivo logístico teniendo en cuenta la estructura jerárquica de los datos (nivel 1: estudiante y nivel 2: centro escolar) propia de los datos educativos.

Muestra

La presente investigación utilizó la base de datos de PISA 2018 proporcionada por la OCDE. En esta última edición, han participado un total de 79 países y aproximadamente 600000 estudiantes de 15 a 16 años. La muestra del estudio está formada por 35943 estudiantes españoles y 1089 centro educativos. La edad media de estos sujetos fue 15.836 años (DT = 0.288), de los cuales el 49.957% son mujeres y el 50.043% son hombres. En la fase inicial de machine learning, se empleó la muestra completa. Sin embargo, en la etapa siguiente de regresión logística multinivel, se excluyeron centros con 20 o menos estudiantes, de tal manera que la muestra fue de 34411 estudiantes provenientes de 976 centros escolares.

Variables

La variable respuesta del estudio es la Comprensión Lectora en PISA 2018, la cual se ha dicotomizado (0 = medio y bajo rendimiento y 1 = alto rendimiento) (ver Tabla 1). Se considera que un estudiante obtiene un alto rendimiento cuando se encuentra en el nivel 5 o 6 en la competencia lectora de PISA 2018 (OCDE, 2019).

Tabla 1

Número de estudiantes españoles con un alto rendimiento en la competencia lectora en PISA 2018

Alto rendimiento lector	N	%
Si	1797	5
No	34146	95

Fuente. Elaboración propia a partir de la OECD (2019).

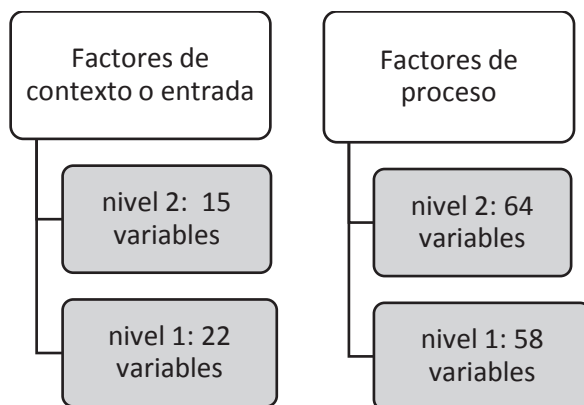
Con relación a las variables predictoras, la presente investigación está conformada por un total de 159 predictores (ver anexo 1), los cuales se encuentran agrupados en factores de contexto o entrada y proceso. Además, estas a su vez se encuentran agrupadas en dos niveles: estudiante (nivel 1) y centro escolar (nivel 2) (ver Figura 2). Cabe destacar, que los predictores están formados por los índices complejos, pero también se han incluido algunas variables directas del cuestionario del estudiante, centro y recorrido educativo. A la hora de seleccionar los factores

contextuales o entrada y de proceso, se han seguido investigaciones previas realizadas por Gamazo et al. (2018), Hu et al. (2021) y Martínez-Abad et al. (2020), las cuales abordan la eficacia escolar en la evaluación internacional de PISA.

En la siguiente figura se muestra el esquema de las variables introducidas en el modelo predictivo.

Figura 2

Esquema de las variables predictoras del estudio



Procedimiento y análisis de datos

Para dar respuesta al primer objetivo, que consiste en identificar los factores de eficacia escolar que tienen una mayor relación con el alto rendimiento lector, se utilizó Random Forest de clasificación. Se ha seleccionado este algoritmo supervisado de machine learning, ya que es uno de los más precisos a la hora de jerarquizar las variables predictoras más importantes (Sterne, 2018). La implementación de este algoritmo demanda un detallado y meticuloso preprocesamiento de los datos previo a su ejecución (Kassambara, 2018). Para ello, en primer lugar, se imputaron los valores perdidos de la base de datos del estudiante y del centro educativo. Se seleccionó el método de imputación múltiple mediante ecuaciones encadenadas, ya que es el más adecuado a la hora de obtener unas estimaciones precisas (Sterne et al., 2009). De esta manera se redujo el conjunto inicial de 159 variables predictoras a 137 tras descartar las que tenían más de un 20% de datos ausentes, tal y como manifiesta Medina y Galván (2007). Se hicieron excepciones con las variables LMMINS (tiempo de aprendizaje en lectura), MMINS (tiempo de aprendizaje en matemáticas) y SMINS (tiempo de aprendizaje en ciencias), las cuales presentaron porcentajes de

ausencia del 24%, 24% y 25%, respectivamente. Dado que el tiempo de aprendizaje juega un rol crucial como una variable de proceso (Martínez-Abad et al., 2020; Hu et al., 2021). En segundo lugar, se procedió a realizar la división de los datos de manera aleatoria en dos conjuntos; el primer conjunto es el de entrenamiento (60% de los datos) con el que se ajustó el modelo y posteriormente con la muestra de validación (40%) se evaluó el rendimiento del modelo (Kassambara, 2018). En tercer lugar, se estandarizaron las variables continuas con el propósito de evitar que los predictores con una mayor magnitud tuvieran una mayor repercusión en el modelo (Sterne et al., 2009). Por último, se analizó la varianza de los predictores, ya que es necesario eliminar las variables predictoras con varianza cero o próxima a cero. No se eliminó ninguna, ya que todas aportan información, es decir tienen variabilidad.

Para obtener el mejor modelo se optimizaron los hiperparámetros mediante la validación cruzada de 10 pliegues. Esto consistió en dividir el conjunto de entrenamiento en subconjuntos del mismo tamaño denominados capas. En la primera iteración el modelo se ajusta con todas las observaciones excepto el primer pliegue que sirve para predecir. En la segunda el modelo se entrena con todas las observaciones excepto el segundo pliegue que sirve para predecir y así sucesivamente hasta la décima iteración (Sarkar et al., 2018).

Se utilizó la técnica de Random Forest en cada uno de los 10 valores plausibles (VP) y se seleccionó aquel con una mayor área bajo la curva (AUC), ya que esta métrica es la más adecuada cuando los niveles de la variable criterio están desbalanceados (Bonaccorso, 2017). El VP más preciso fue el nueve (0.827), por lo que se utilizó este VP a la hora de reportar las variables más importantes. La gran limitación de esta técnica es que no hay un punto de corte a la hora de seleccionar los predictores que tienen una mayor influencia en la variable respuesta (Sarkar et al., 2018), por ello, siguiendo el estudio de Gorostiaga y Rojo-Álvarez (2016) quienes recomienda que para seleccionar el número óptimo de variables es necesario evaluar varios conjuntos de variables (20 y 30) y seleccionar aquel conjunto con un mejor rendimiento, se decidió examinar de los 137 predictores los siguientes conjuntos de variables: 15, 20, 25, 30 y 35. De esta manera, se aporta una mayor precisión a la hora de indicar qué conjunto de variables es el más óptimo. El método seleccionado a la hora de ordenar las variables es el de importancia relativa.

El segundo objetivo se basó en determinar la contribución relativa de los factores del contexto o entrada y proceso que más inciden en la probabilidad de que un estudiante obtenga un alto rendimiento en la competencia lectora teniendo en cuenta la estructura jerárquica de los datos (nivel 1: estudiante y nivel 2: centro escolar). Debido a que la variable criterio es de naturaleza cualitativa nominal dicotómica, la varianza aleatoria fue significativa en el nivel 2 (centro escolar) y el Coeficiente de Correlación Intraclase fue superior al 10% (Lee, 2000) se realizó la técnica de regresión logística binaria multinivel. Antes de llevar a cabo los modelos

se comprobó el supuesto de multicolinealidad a través de dos pasos. En primer lugar, se correlacionaron las variables mediante el coeficiente de correlación de Spearman, ya que ninguna de las variables cumple con el supuesto de normalidad. El trabajo realizado por Kassambara (2018) indica que se considera una alta intensidad de relación aquellos valores superiores a 0.7. En la tabla 2, se aprecia que existen dos asociaciones con una correlación de magnitud alta. Concretamente, se eliminaron las variables *índice posición ocupacional más alta de los padres*, y *bienes del hogar* ya que tienen una menor importancia relativa que el *índice de situación socioeconómico y cultural*.

Tabla 2

Correlación de Spearman

	Índice de situación socioeconómico y cultural
Índice posición ocupacional más alta de los padres	0.815***
Bienes del hogar	0.704***

Nota. (*) Diferencia significativa con un $\alpha = .05$ (**) significativa con un $\alpha = .01$; (***) significativa con un $\alpha = .001$.

En segundo lugar, se comprobó que el Valor de la Inflación de la Varianza fuera inferior a 10 y como en todos los casos es así, se llegó a la conclusión de que se cumple el supuesto de multicolinealidad.

A continuación, se eliminaron los centros educativos con menos de 20 sujetos para realizar un correcto análisis multinivel (Gamazo et al., 2018). Concretamente, se han excluido un total de 1532 sujetos pertenecientes a 133 centros, por lo que la muestra final fue de 34411 estudiantes. Se generaron cuatro modelos con las variables de contexto o entrada y proceso más importantes seleccionadas por el algoritmo de machine learning. El modelo 0 o nulo sin predictores; el modelo 1 que está formado por las variables de contexto o entrada; el modelo 2 compuesto por las variables de proceso; por último, el modelo 3 formado por todas las variables. De esta manera, se analiza la contribución de las variables de proceso una vez controlados los factores de contexto o entrada. A la hora de interpretar los coeficientes se utilizó la *odds ratio* (razón de probabilidades) (Lee, 2000), referido a la probabilidad de que un estudiante tenga un alto rendimiento en comprensión lectora, y la probabilidad de que no lo tenga. También se estimó el Porcentaje de Varianza Explicado (PVE), el cual señala la cantidad de varianza que explica el modelo (ecuación 1).

$$PVE = \left(\frac{\sigma_{\text{Modelo nulo}}^2 - \sigma_{\text{Modelo final}}^2}{\sigma_{\text{Modelo nulo}}^2} \right) * 100 \quad \text{Ecuación 1}$$

Finalmente, para comprobar el ajuste de los modelos se utilizaron los índices AIC, BIC y Deviance, también, con el fin de comparar modelos anidados, se calculó la significatividad de este estadístico de reducción y se estimó el porcentaje de varianza reducido (R^2) (Cameron & Windmeijer, 1997). Las variables fueron introducidas en el modelo según el orden obtenido con el algoritmo de machine learning, tal y como se llevó a cabo en los trabajos realizados por Arroyo et al. (2019) y Constante-Amores et al. (2021).

Todos los análisis se realizaron con el software estadístico R versión 4.3.3. Se utilizaron los paquetes mice y lme4 para llevar a cabo el método de imputación múltiple mediante ecuaciones encadenadas y regresión logística binaria multinivel, respectivamente. El modelo de Random Forest se ha realizado con la librería de machine learning H2o, la cual está escrita en Java, pero se puede utilizar con el lenguaje de programación R.

RESULTADOS

A continuación, se presentan los resultados en función de los objetivos planteados.

Objetivo específico 1: Factores de contexto o entrada más importantes

En la siguiente tabla se puede apreciar que el primer conjunto de predictores formado por 15 variables es el que tiene un mejor rendimiento (AUC = 0.839), es decir, el que mejor representa las características de los estudiantes con un alto rendimiento en la competencia lectora.

Tabla 3

Rendimiento del modelo en la muestra de validación de los cinco conjuntos de características candidatas mediante bosques aleatorios

	AUC	Precisión
15 variables	0.839	0.952
20 variables	0.837	0.947
25 variables	0.830	0.950
30 variables	0.830	0.945
35 variables	0.831	0.949

En la tabla 4 se observan los 15 predictores que más se vinculan con el alto rendimiento en comprensión lectora, de las 139 variables de contexto o entrada y proceso. De manera general se aprecia que todas estas variables están relacionadas con las características del alumnado. En relación con los factores de contexto y proceso, las variables más relevantes son: *índice de situación económica, social y cultural, ocupación del padre e índice de la posición ocupacional más alta padres*. Con respecto a las variables de proceso, las más importantes son el *disfrute de la lectura, metacognición: evaluar credibilidad y autoconcepto en lectura: percepción de competencia*.

Tabla 4

Variables más importantes asociadas al alto rendimiento en comprensión lectora

		Importancia relativa (%)
Contexto o entrada	6º Índice de situación económica, social y cultural	27.432
	8º Ocupación del padre	20.300
	10º Índice de la posición ocupacional más alta padres	13.569
	11º Bienes del hogar	10.736
	12º Ocupación de la madre	10.506
	13º Posesiones culturales en el hogar	10.497
Proceso	1º Me gusta o disfruto con la lectura	100.000
	2º Metacognición: evaluar la credibilidad	86.520
	3º Autoconcepto lector: percepción de competencia	61.827
	4º Autoconcepto lector: percepción de dificultad	39.826
	5º Tiempo de aprendizaje en ciencias (minutos por semana)	32.459
	7º Metacognición: resumir	23.747
	9º Situación laboral esperado por el alumnado	19.379
	14º Uso de las TIC fuera de la escuela (para actividades escolares)	7.167
	15º Las TIC como tema de interacción social	4.127

Objetivo específico 2: probabilidad de obtener un alto rendimiento lector

Una vez identificadas las variables más importantes se procedió a realizar el análisis logístico binario multinivel. La tabla 5 muestra los resultados de los modelos predictivos. En el modelo 1 compuesto por las variables de contexto o

de entrada, se aprecia que todas han resultado estadísticamente significativas con direccionalidad positiva y explican aproximadamente el 41% de la varianza de la variable criterio. El efecto más importante se da en el predictor *me gusta o disfruto con la lectura*. Específicamente, por cada punto más en este predictor, la probabilidad de obtener un alto rendimiento lector aumenta en un 100%. El modelo 2, formado únicamente por las variables de proceso, al igual que en el anterior modelo, todas resultan estadísticamente significativas, explicando el 38% del alto rendimiento en la competencia lectora. El efecto más alto se da en la variable *metacognición: evaluar la credibilidad*. Concretamente, por cada punto más en este predictor, la probabilidad de obtener un alto rendimiento en la competencia lectora aumenta en un 110%. En el modelo 3 se añaden todas las variables resultando las mismas estadísticamente significativas. Además, presentan una dirección positiva, excepto *autoconcepto lector: percepción de competencia, uso de las TIC fuera de la escuela (para actividades escolares)* y las *TIC como tema de interacción social*. La variable que tiene un mayor efecto es *metacognición: evaluar la credibilidad*. Una vez controlado las variables de contexto o entrada los factores de proceso, se incrementa el porcentaje de varianza explicada alrededor de un 13%, respecto al modelo 1, de esta manera se obtiene una varianza total explicada del 54.356 %

Por último, en lo que concierne al ajuste de los modelos, se observa que el modelo con un menor AIC y BIC es el modelo compuesto por todos los predictores (modelo 3), donde se consigue una reducción significativa de la varianza equivalente a un R^2 del 11%, respecto al modelo 1. También fue significativa la reducción de la varianza del modelo 1 respecto al modelo 0 (12%), y del modelo 2 respecto al nulo (20%).

Tabla 5

Estimación de efectos fijos con errores estándar robustos mediante regresión logística binaria multinivel (Odds ratio)

	Modelo 0 (nulo)	Modelo 1 (contexto o entrada)	Modelo 2 (proceso)	Modelo 3 (todas)
Intercepto	0,041 ***	0,015 ***	0,002***	0,001***
Índice de situación económica, social y cultural		1.245 ***		1.152***
Ocupación del padre		1.008 ***		1.007***
Ocupación de la madre		1.007 ***		1.005***
Posesiones culturales en el hogar		1.195 ***		1.145***
Disfrute de la lectura			2.000 ***	1.369***

	Modelo 0 (nulo)	Modelo 1 (contexto o entrada)	Modelo 2 (proceso)	Modelo 3 (todas)
Metacognición: evaluar la credibilidad			2.100***	1.989***
Autoconcepto lector: percepción de competencia			1.875***	1.436***
Autoconcepto lector: percepción de dificultad			0.782***	0.790***
Tiempo de aprendizaje en ciencias (minutos por semana)			1.002***	1.001***
Metacognición: resumir			1.777***	1.358***
Ocupación laboral esperada por el alumnado			1.027***	1.022***
Uso de las TIC fuera de la escuela (para actividades escolares)			0.852***	0.809***
Las TIC como tema de interacción social			0.788***	0.819***
Varianza	0.528	0.312	0.326	0.241
PVE (%)		40.909	38.258	54.356
AIC	14881	13045	11884	11509
BIC	14895	13104	11969	11636
Loglik	-7563	-6515	-5932	-5739
Deviance	14850	13031	11864	11479
X^2		1818	2985	1552
Pr(>Chisq)		0.000	0.000	0.000

Nota. (*) Diferencia significativa con un $\alpha = .05$ (**) significativa con un $\alpha = .01$; (***) significativa con un $\alpha = .001$.

DISCUSIÓN Y CONCLUSIONES

En este trabajo se ha profundizado en el estudio de los factores (contexto o entrada y proceso) de eficacia escolar asociados al alto rendimiento en comprensión lectora en el marco de PISA 2018, el cual es uno de los grandes problemas del sistema educativo.

Nuestros resultados identifican 15 variables con una mayor influencia en el alto rendimiento lector, de entre las cuales la más importante es la variable de proceso

referida al gusto por la lectura. Estos resultados también se observan en los trabajos realizados por Akande y Oyedapo (2018), Franco et al. (2016) y Molina (2020) que indican la importancia de este predictor en el éxito académico lector. En este sentido, existe una amplia literatura (Gil, 2011) que indica la relevancia que tiene la familia a la hora de fomentar el placer por la lectura. Además, la investigación de Butlen (2005) muestra que el centro escolar también desempeña un papel importante. Por ello, desde la escuela se tienen que llevar a cabo medidas educativas para incrementar en el alumnado el disfrute por la lectura (Dezcallar et al., 2014), siendo fundamental trabajar en los centros el hábito lector de los estudiantes.

Se observa también que los cinco predictores más importantes pertenecen a los factores de proceso (placer por la lectura, metacognición: evaluar la credibilidad, autoconcepto lector entendido como percepción de competencia y percepción de dificultad y tiempo de aprendizaje en ciencias -minutos por semana-). Por lo que, a la vista de estos resultados parece que estos factores juegan un papel muy relevante en el alto rendimiento en comprensión lectora. Esto dista de una amplia literatura que manifiesta que el alto desempeño lector es explicado fundamentalmente por factores de contexto o entrada (Cordero et al., 2013; Franco et al., 2016). Si bien coincide con el trabajo de Martínez-Abad et al. (2020), en el que las variables más importantes están relacionadas con el alumnado.

Asimismo, hay que destacar que en nuestro estudio el sexo del estudiante y la condición de inmigrante no son variables importantes en el alto rendimiento lector, a diferencia de otros estudios (de Frutos & Santaren, 2020; Cordero et al., 2013), incluso Hu et al. (2021) demuestran que estos predictores son importantes a la hora de discriminar y diferenciar entre los estudiantes de alto y bajo rendimiento científico en PISA 2015.

Con respecto al segundo objetivo, basado en determinar la contribución relativa de los factores que más inciden en la probabilidad de que un estudiante obtenga un alto desempeño en la competencia lectora, se puede apreciar la importancia que tienen, tanto los factores de contexto o entrada como de proceso en el rendimiento lector; puesto que cada conjunto de predictores explica un porcentaje varianza muy similar 39% y 36%, respectivamente. El modelo final explica aproximadamente el 52% de la variabilidad de la variable criterio. Estos resultados complementan las conclusiones de Coleman (1966) que mostró la preponderancia de factores de contexto o entrada como nivel socioeconómico y etnia.

En el modelo final, el predictor que tiene un mayor efecto es la metacognición (odds ratio = 1.989), referido a la capacidad del sujeto para evaluar la calidad y credibilidad de un texto (por ejemplo, si la información es válida, precisa e imparcial, etc.) incorporado por primera vez en PISA 2018 (OCDE, 2019). Este predictor tiene aproximadamente el doble de efecto que ciertas variables de contexto o entrada que han sido de gran importancia en la literatura científica como son el índice de

situación económica, social y cultural, ocupación del padre, ocupación de la madre y posesiones culturales del hogar (Barrera et al., 2019). De hecho, la otra variable relacionada con la metacognición (resumir), también tiene un mayor efecto que este conjunto de predictores.

En este sentido, destaca la gran importancia que tiene la metacognición en el alto rendimiento lector frente a otras variables de corte más socioeconómico y se sigue la línea de Qi (2021) y Wu et al. (2020) que señalan el importante papel que juegan las variables metacognitivas en el alto rendimiento lector, por lo que es fundamental que desde las escuelas se trabaje la capacidad metacognitiva.

Asimismo, al igual que en las investigaciones de Franco et al. (2016) y Herrera et al. (2017), hay que destacar la relevancia que tiene el autoconcepto lector (percepción de competencias y percepción de dificultad) en las altas puntuaciones de lectura en PISA 2018. También juegan un papel muy importante las expectativas educativas académicas y profesionales del alumnado sobre la competencia lectora, tal y como ocurría en el estudio de Hu et al. (2021).

Desde el punto de vista metodológico, esta investigación no indica únicamente las variables más importantes, sino que también lleva a cabo un modelo predictivo (regresión logística binaria multinivel) abordando de manera exhaustiva este fenómeno educativo para un grupo específico que son los estudiantes con alto rendimiento frente al resto. También este estudio supone un avance metodológico en el fenómeno de los factores de eficacia escolar, ya que se emplea por primera vez un enfoque de machine learning en el contexto español. Este es un enfoque complementario a los trabajos que utilizan técnicas de minería de datos (Martínez-Abad et al., 2020) y regresión logística binaria (Gamazo et al., 2018).

Con el objetivo deseable para el sistema educativo español de aumentar el escaso número de estudiantes de alto rendimiento, podrían desarrollarse políticas educativas que fomenten el placer por la lectura y el aprendizaje de herramientas meta-cognitivas en el ámbito de la lectura, bien conocidas y necesarias, como la evaluación de la credibilidad de los textos y la capacidad de resumir, acreditadas por los resultados de esta investigación. Además, desde los centros educativos es necesario que se trabaje en la mejora del autoconcepto, ya que el mismo repercute en las expectativas académicas de los estudiantes y, por ende, en el rendimiento de diferentes materias (Carrillo et al., 2022).

Como prospectiva de investigación, sería necesario determinar los factores (contexto o entrada y proceso) asociados al alto rendimiento lector en otros países y contextos, como por ejemplo algunos del sur de Europa (Portugal, Italia, Francia, Chipre, Grecia y Malta). Respecto a las limitaciones del estudio, hay que destacar que los resultados obtenidos no pueden interpretarse en términos de causalidad, para ello, sería necesario realizar modelos de ecuaciones estructurales o plantear estudios con carácter experimental. Además, la evaluación PISA ha sido revisada y

criticada metodológicamente, como muestra el trabajo de Fernández-Cano (2016) aludiendo, entre otras cuestiones, a los problemas de validez de los instrumentos de medida y la forma de estimar las puntuaciones de la escala de rendimiento. Otro problema son los cuestionarios utilizados para recoger información de contexto (Jornet, 2016), que carecen de una teoría clara para elaborar los diferentes constructos medidos y la falta de información sobre sus características psicométricas.

REFERENCIAS BIBLIOGRÁFICAS

- Akande, S., & Oyedapo, R. (2018). Developing the reading habits of secondary school students in Nigeria: the way forward. *International Journal of Library Science*, 7(1), 15-20.
- Arroyo, D., Constante-Amores, I. A., & Asensio, I. (2019). La repetición de curso a debate: un estudio empírico a partir de PISA 2015. *Educación XX1*, 22(2), 69-92. <https://doi.org/10.5944/educxx1.22479>
- Artelt, C., & Schneider, W. (2015). Cross-country generalizability of the role of metacognitive knowledge in students' strategy use and reading competence. *Teachers College Record*, 117(1), 1–32. <https://bit.ly/4akHyKl>
- Asensio-Muñoz, I., Carpintero, E., Expósito, E., & López-Martín, E. (2018). ¿Cuánto oro hay entre la arena? Minería de datos con los resultados de España en PISA 2015. *Revista de Educación*, 270, 225-245
- Avendaño, I., & Martínez, D. (2013). Competencia lectora y el uso de las nuevas tecnologías de la información y comunicación. *Revista Escenarios*, 11(1), 7-22.
- Barrera, J. E., Polanco, J. G. C., & Acosta, J. D. (2019). Comprensión lectora de estudiantes universitarios. Factores asociados y mecanismos de acción. *Revista Venezolana de Gerencia*, 24(87), 874-889.
- Bonaccorso, G. (2017). *Machine learning algorithms*. Packt Publishing Ltd.
- Butlen, M. (2005). Paradoja de la lectura escolar. *Revista de Educación*, núm. Extraordinario, 139-151.
- Cameron, A. C., & Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, 77(2), 329-342.
- Campos, I. O., & Arantegui, M. (2022). Exploración de la mediación parental en el uso de las TIC y su correlación con la comprensión lectora del alumnado preadolescente. *Lenguaje y Textos*, 55, 43-54. <https://doi.org/10.4995/lyt.2022.15948>
- Carrillo-López, P. J., Constante-Amores, A., Arroyo-Resino, D., & Sánchez-Munilla, M. (2022). Self-concept and academic achievement in primary school: a predictive study. *International Journal of Education in Mathematics, Science and Technology*, 10(4), 1057-107

- Constante-Amores, A., Florenciano Martínez, E., Navarro, E., & Fernández-Mellizo, M. (2021). Factores asociados al abandono universitario. *Educación XX1*, 24(1), 17-44.
- Coleman, J. (1966). *Equality of educational opportunity*. US Government Printing Office.
- Cordero, J. M., Crespo, E., & Pedraja, F. (2013). Rendimiento educativo y determinantes según PISA: una revisión de la literatura en España. *Revista de Educación*, 362, 273-297
- Creemers, B. P., Peters, T., & Reynolds, D. (2022). *School effectiveness and school improvement*. Routledge
- de Frutos, S. F., & Santaren, V. R. (2020). El papel del sexo en comprensión lectora. Evidencias desde PISA y PIRLS. *Revista de Investigación en Educación*, 18(2), 99-117.
- de la Orden, A., Asensio, I., Carballo, R., Fernández Díaz, J., Fuentes, A., García Ramos, J. M., & Guardia, S. (1997). Desarrollo y validación de un modelo de calidad universitaria como base para su evaluación. *RELIEVE - Revista Electrónica De Investigación Y Evaluación Educativa*, 3(1).
- Dezcallar, T., Clariana, M., Cladellas, R., Badia, M., & Gotzens, C. (2014). La lectura por placer: su incidencia en el rendimiento académico, las horas de televisión y las horas de videojuegos. *Ocnos: Revista de Estudios sobre Lectura*, (12), 107-116.
- Franco, M., Cárdenas, R., & Santrich, E. (2016). Factores asociados a la comprensión lectora en estudiantes de noveno grado de Barranquilla. *Psicogente*, 19(36), 296-310. <https://doi.org/10.17081/psico.19.36.1299>
- Gamazo, A., Martínez-Abad, F., Olmos-Miguelañez S., & Rodríguez-Conde M. J. (2018). Evaluación de los factores relacionados con la eficacia escolar en PISA. Un análisis multinivel. *Revista de Educación*, 379, 56-84.
- García M. A., Arévalo M. A., & Hernández C. A. (2018). La comprensión lectora y el rendimiento escolar. *Cuadernos de Lingüística Hispánica*, (32), 155- 174.
- Garrido, R., Gallo-Rivera, M. T., & Martínez-Gautier, D. (2020). ¿Cuáles son y cómo operan los determinantes del fracaso escolar? Replanteando las políticas públicas para el caso de España y sus regiones. *Revista Internacional de Ciencias del Estado y de Gobierno*, 1(4), 509-540.
- Gaviria, J. L. (2004). La situación española: el rendimiento de los estudiantes. En G. Haug, J. L. Gaviria, C. Lomas, M.D. de Prada, & D. Gil (Eds.), *El rendimiento de los estudiantes al final de la educación obligatoria: objetivos europeos y situación española* (pp.18- 83). Santillana.
- Gil, J. (2011). Hábitos lectores y competencias básicas en el alumnado de educación secundaria obligatoria. *Educación XX1*, 14(1), 117-134.

- Gómez, I. M., García, J. A., Vila, J. O., Elosúa, M. R., & Rodríguez, R. (2014). The dual processes hypothesis in mathematics performance: beliefs, cognitive reflection, working memory and reasoning. *Learning and Individual Differences*, (29), 67-73.
- Gorostiaga, A., & Rojo-Álvarez, J. L. (2016). On the use of conventional and statistical-learning techniques for the analysis of PISA results in Spain. *Neurocomputing*, 171, 625-637.
- Herrera, J. C., Treviño, A., & Navarrete, G. (2017). *Factores de riesgo asociados a falta de competencia para la comprensión lectora en niños de primaria* [Comunicación]. X Congreso Nacional de Investigación Educativa.
- Hu, J., Peng, Y., & Ma, H. (2021). Examining the contextual factors of science effectiveness: a machine learning-based approach. *School Effectiveness and School Improvement*, 33, 21–50. <https://doi.org/10.1080/09243453.2021.1929346>
- Kassambara, A. (2018). *Machine learning essentials: Practical guide in R*. Sthda.
- Kyriakides, L., Charalambous, E., Creemers, B. P. M., Antoniou, P., Devine, D., Papastilianou, D., & Fahie, D. (2019). Using the dynamic approach to school improvement to promote quality and equity in education: a European study. *Educational Assessment, Evaluation and Accountability*, 31(1), 121–149. <https://doi.org/10.1007/s11092-018-9289-1>
- Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: the case of school effects. *Educational Psychologist*, 35(2), 125-141.
- Linnakyla, P., Malin, A., & Taube, K. (2004). Factors behind low reading literacy achievement. *Scandinavian Journal of Educational Research*, 48(3), 231-249.
- López-González, E., Navarro, E., García San Pedro, M. J., Lizasoain, L., & Tourón, J. (2021). Estudio de la eficacia escolar en centros educativos de primaria mediante el uso de modelos jerárquicos lineales. *Bordón*, 73(1), 59-80. <https://doi.org/10.13042/Bordon.2021.80530>
- Martínez-Abad, F., Gamazo, A., & Rodríguez-Conde, M. J. (2020). Educational data mining: identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation*, 66, Artículo 100875.
- Medina, F., & Galván, M. (2007). *Imputación de datos: teoría y práctica*. Cepal.
- Molina, I. (2020). Comprensión lectora y rendimiento escolar. *Revista Boletín Redipe*, 9(1), 121-131.
- Murillo, F. J. (Coord.) (2007). *Investigación iberoamericana sobre eficacia escolar*. Convenio Andrés Bello.
- Organización para la Cooperación y el Desarrollo Económicos. (2019). *PISA 2018 assessment and analytical framework*. OECD publishing.
- Qi, X. (2021). Effects of self-regulated learning on student's reading literacy: evidence from Shanghai. *Frontiers in Psychology*, 11, Artículo 555849.

- Rivera, J. (2013). Punto de encuentro entre los jóvenes y la lectura: estrategia constructivista para fortalecer su comportamiento lector. *Simbiosis Estudiantil*, 7(1).
- Sarkar, D., Bali, R., & Sharma, T. (2018). Practical machine learning with Python. A problem-solvers guide to building real-world intelligent systems. Apress.
- Scheerens, J. (1991). Process indicators of school functioning: a selection based on the research literature on school effectiveness. *Studies in Educational Evaluation*, 17(2), 371–403. [https://doi.org/10.1016/S0191-491X\(05\)80091-4](https://doi.org/10.1016/S0191-491X(05)80091-4).
- Scheerens, J., & Creemers, B. P. (2022). *School effectiveness and school improvement*. Routledge
- Scheerens, J., Witziers, B., & Steen, R. (2013). A meta-analysis of school effectiveness studies. *Revista de Educación*, 361, 619-645. <https://doi.org/10.4438/1988-592X-RE-2013-361-235>
- Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 338, 23-93
- Viramontes, E., Amparán, A., & Núñez, L. D. (2019). Comprensión lectora y el rendimiento académico en educación primaria. *Investigaciones sobre Lectura*, 12, 65-82.
- Wu, Y. J., Carstensen, C. H., & Lee, J. (2020). A new perspective on memorization practices among East Asian students based on PISA 2012. *Educational Psychology*, 40(5), 643-662.

ANEXO 1

Tabla 1

Variables de contexto o entrada

	SC001Q01TA	¿Cuál de las siguientes definiciones describe mejor la comunidad en la que se encuentra su escuela?
	SCHLTYPE	Titularidad del centro escolar
	STRATIO	Ratio de alumno por profesor
	SCHSIZE	Tamaño de la escuela
	CLSIZE	Tamaño de la clase
	EDUSHORT	Escasez de material educativo
	STAFFSHORT	Escasez de personal educativo
	RATCMP1	Número de ordenadores disponibles por estudiante
	RATCMP2	Proporción de ordenadores disponibles que están conectados a Internet
Nivel 2: centro	TOTAT	Número total de profesores del centro escolar
	PROATCE	Índice de proporción de profesores certificados
	PROAT5AB	Índice de proporción de profesores con Licenciatura
	PROAT5AM	Índice de proporción de profesores con Máster
	PROAT6	Índice de proporción de profesores con Doctorado
	SC002Q01TA	Número de chicos
	SC002Q02TA	Número de chicas
	SC048Q03NA	Porcentaje de alumnos de 15 años que proceden de hogares socioeconómicamente desfavorecidos
	SC048Q01NA	Estudiantes cuya lengua materna es diferente del castellano.
	ST004D01T	Género
	ST003D02T	Mes de nacimiento.
Nivel 1: estudiante	SCCHANGE	Número de cambios del colegio
	PAREDINT	Índice de estudios superiores de los padres
	BMMJ1	ISEI de la madre
	BFMJ2	ISEI del padre

	IMMIG	Índice de situación de inmigración
	DURECEC	Duración de la educación y atención a la primera infancia.
	REPEAT	Repetición de curso
	CHANGE	Números de cambios en la biografía educativa
	ESCS	Índice de situación socioeconómica y cultural
	ICTHOME	TIC disponible en el hogar
	HOMEPOS	Posesiones del hogar
	CULTPOSS	Cultural possessions at home
Nivel 1: estudiante	WEALTH	Riqueza familiar
	MISCED	Educación de la madre
	FISCED	Educación del padre
	HISCED	Educación más alta de los padres
	PARED	Índice de la educación más alta de los padres en años de escolarización
	HEDRES	Recursos educativos del hogar
	ICTRES	Recursos TIC del hogar
	HISEI	Índice de la situación profesional más alta de los padres

Table 2
Variables de proceso del nivel 2

Dificultad de aprendizaje	STUBEHA	Comportamiento de los alumnos que dificultan el aprendizaje
	TEACHBEHA	Comportamiento del profesor que dificulta el aprendizaje
Participación familiar	SC064Q01TA	Proporción de padres: discutieron el progreso de su hijo con un profesor por iniciativa propia.
	SC064Q02TA	Proporción de padres: hablaron de los progresos de su hijo por iniciativa de uno de sus profesores.
	SC064Q03TA	Proporción de padres: participaron en el gobierno escolar local (por ejemplo, en el consejo de padres o en el comité de gestión escolar).
	SC064Q04TA	Proporción de padres: participaron como voluntarios en actividades físicas o extraescolares [...].
Nivel 2: centro escolar	SC037Q01TA	Garantía de calidad en la escuela: evaluación interna/autoevaluación.
	SC037Q02TA	Garantía de calidad en la escuela: evaluación externa.
	SC037Q03TA	Garantía de calidad en el centro educativo: especificación escrita del perfil curricular y de los objetivos educativos del centro.
	SC037Q04TA	Garantía de calidad en el centro educativo: especificación por escrito de los estándares de rendimiento de los alumnos.
	SC037Q05NA	Garantía de calidad en el centro educativo: registro sistemático de datos como [...] la asistencia y el desarrollo profesional.
	SC037Q06NA	Garantía de calidad en la escuela: registro sistemático de los resultados de las pruebas de los alumnos y de las tasas de graduación.
	SC037Q07TA	Garantía de calidad en la escuela: Recabar la opinión por escrito de los alumnos (por ejemplo, sobre las clases, los profesores o los recursos).
	SC037Q08TA	Garantía de calidad en la escuela: la tutoría de los profesores.
	SC037Q09TA	Garantía de calidad en la escuela: consultas periódicas para la mejora del centro escolar [...] durante un período mínimo de seis meses.
	SC037Q010NA	Garantía de calidad en la escuela: aplicación de una política normalizada para las asignaturas de lectura [...].

SC154Q01HA	Uso de las evaluaciones de los alumnos por parte de la escuela: para guiar el aprendizaje de los alumnos.
SC154Q02WA	Uso de las evaluaciones de los alumnos por parte de la escuela: para informar a los padres sobre el progreso de sus hijos.
SC154Q03WA	Uso de las evaluaciones de los alumnos por parte de la escuela: para tomar decisiones sobre la permanencia o promoción de los alumnos.
SC154Q04WA	Uso de las evaluaciones de los alumnos por parte de la escuela: para agrupar a los estudiantes con fines educativos.
SC154Q05WA	Uso de las evaluaciones de los alumnos por parte de la escuela: para comparar el rendimiento de la escuela con el del <distrito o el país>.
SC154Q06WA	Uso de las evaluaciones de los estudiantes por parte de la escuela: para monitorear el progreso de la escuela de un año a otro.
SC154Q07WA	Uso de las evaluaciones de los alumnos por parte de la escuela: para juzgar la eficacia de los profesores.
SC154Q08WA	Utilización de las evaluaciones de los alumnos por parte de la escuela: para identificar aspectos de la enseñanza o del plan de estudios que podrían mejorarse.
SC154Q09HA	Utilización de las evaluaciones de los alumnos por parte de la escuela: para adaptar la enseñanza a las necesidades de los alumnos.
SC154Q10WA	Utilización de las evaluaciones de los alumnos por parte de la escuela: para comparar la escuela con otras escuelas.
SC154Q011HA	Utilización de las evaluaciones de los alumnos por parte del centro: para otorgar certificados a los alumnos.
SC165Q09HA	En nuestro centro educativo, se anima a los alumnos a comunicarse con personas de otras culturas a través de [...].
SC165Q10HA	En nuestro centro educativo adopta diferentes enfoques para educar a los alumnos sobre las diferencias culturales [...]
SC156Q01HA	En la escuela: su propia declaración escrita sobre el uso de dispositivos digitales.
SC156Q02HA	En la escuela: su propia declaración escrita específicamente sobre el uso de dispositivos digitales con fines pedagógicos.

Evaluación de los
estudiantes

Nivel 2: centro
escolar

Autonomía del centro

	SC156Q04HA	En la escuela: un programa de uso de dispositivos digitales para la enseñanza y el aprendizaje en asignaturas específicas.
	SC156Q05HA	En la escuela: conversaciones periódicas con el personal docente sobre el uso de dispositivos digitales con fines pedagógicos.
	SC156Q06HA	En la escuela: un programa específico para preparar a los alumnos para un comportamiento responsable en Internet.
	SC156Q07HA	En la escuela: una política específica sobre el uso de las redes sociales (<Facebook>, etc.) en la enseñanza y el aprendizaje.
	SC156Q08HA	En la escuela: un programa específico para promover la colaboración en el uso de dispositivos digitales entre los profesores.
	SC156Q09HA	En la escuela: tiempo programado para que los profesores se reúnan para compartir, evaluar o desarrollar materiales didácticos y [...]
	SC155Q01HA	Capacidad de la escuela mediante dispositivos digitales: El número de dispositivos digitales conectados a Internet es suficiente
	SC155Q02HA	Capacidad del centro escolar mediante dispositivos digitales: El ancho de banda o la velocidad de Internet del centro es suficiente.
	SC155Q03HA	Capacidad del centro para utilizar dispositivos digitales: El número de dispositivos digitales para la enseñanza es suficiente
	SC155Q04HA	Capacidad del centro para utilizar dispositivos digitales: los dispositivos digitales [...] son suficientemente potentes en términos de capacidad informática
	SC155Q05HA	Capacidad del centro para utilizar dispositivos digitales: la disponibilidad de software adecuado es suficiente
	SC155Q06HA	Capacidad del centro para utilizar dispositivos digitales: Los profesores tienen las [...] competencias necesarias para integrar los dispositivos digitales en la enseñanza
	SC155Q07HA	Capacidad del centro para utilizar dispositivos digitales: Los profesores disponen de tiempo suficiente para preparar las clases integrando dispositivos digitales
	SC155Q08HA	Capacidad del centro para utilizar dispositivos digitales: recursos profesionales eficaces para que los profesores aprendan a utilizar [...]
Nivel 2: centro escolar	Autonomía del centro	

SC155Q09HA	Capacidad del centro para utilizar dispositivos digitales: se dispone de una plataforma eficaz de apoyo al aprendizaje en línea
SC155Q10HA	Capacidad del centro para utilizar dispositivos digitales: los profesores reciben incentivos para integrar los dispositivos digitales en [...].
SC155Q11HA	Capacidad del centro para utilizar dispositivos digitales: el centro cuenta con suficiente personal técnico auxiliar cualificado
SC042Q01TA	Política del centro para <grado modal nacional para alumnos de 15 años>: Los alumnos se agrupan por capacidades en clases diferentes.
SC042Q02TA	Política del centro para <grado modal nacional para alumnos de 15 años>: Los alumnos se agrupan por capacidades dentro de sus clases.
SC150Q01IA	Políticas del centro orientadas a la equidad: estos alumnos asisten a clases regulares y reciben períodos adicionales de [...]
SC150Q02IA	Políticas del centro orientadas a la equidad: antes de pasar a las clases ordinarias, [...] programa preparatorio destinado a [...]
SC150Q03IA	Políticas del centro orientadas a la equidad: Antes de pasar a las clases normales, [...] instrucción en materias escolares [...]
SC150Q04IA	Políticas de la escuela orientadas a la equidad: Estos alumnos reciben [...] cantidades de instrucción en su <lengua materna>
SC150Q05IA	Políticas del centro orientadas a la equidad: El tamaño de las clases se reduce para atender las necesidades especiales de estos alumnos.
TMINS	Tiempo de aprendizaje (minutos por semana) en total
MMINS	Tiempo de aprendizaje (minutos por semana) en Matemáticas
LMINS	Tiempo de aprendizaje (minutos por semana) en lengua
SMINS	Tiempo de aprendizaje (minutos por semana) en Ciencias
BSMJ	Situación profesional esperada por el estudiante
ST225Q01HA	¿Espera completarlo? <Nivel 2 de la CINE>
ST225Q02HA	¿Espera completarlo? <Nivel CINE 3B o C>
ST225Q03HA	¿Espera completarlo? <Nivel 3A del CINE>
ST225Q05HA	¿Espera completarlo? <Nivel ISCED 5B>
Nivel 2: centro escolar	Autonomía del centro
Nivel 1: estudiante	

ST225Q06HA	¿Espera completar? <Nivel ISCED 5A o 6>
JOYREAD	Placer por la lectura
UNDREM	Metacognición: comprender y recordar
METASUM	Metacognición: resumir
METASPAM	Metacognición: evaluar la credibilidad
MASTGOAL	Orientación hacia el dominio de los objetivos de rendimiento
DISCLIMA	Disciplinary climate in test language lessons
PERFEED	Retroalimentación percibida
SCREADCOMP	Autoconcepto de la lectura: percepción de competencia
SCREADDIFF	Autoconcepto de la lectura: percepción de la dificultad
PISADIFF	Percepción de la dificultad de la prueba PISA
PERCOMP	Percepción de la competitividad en la escuela
PERCOOP	Percepción de la cooperación en la escuela
ENTUSE	Uso de las TIC fuera de la escuela (ocio)
HOMESCH	Uso de las TIC fuera de la escuela (para actividades escolares)
ICTSCH	TIC disponibles en la escuela
TEACHSUP	Apoyo del profesor en las clases de lengua de prueba
USESCH	Uso de las TIC en la escuela en general
COMPACT	Competencia percibida en TIC
INTICT	Interés por las TIC
AUTICT	Autonomía percibida en relación con el uso de las TIC
SOIAICT	Las TIC como tema de interacción social
ICTCLASS	Uso de las TIC relacionado con la asignatura durante las clases
ICTOUTSIDE	Uso de las TIC relacionado con la asignatura fuera de clase
TEACHINT	Interés percibido del profesor
PASCHPOL	Políticas escolares para la participación de los padres
DIRINS	Enseñanza dirigida por el profesor
STIMREAD	Estimulación del compromiso con la lectura por parte del profesor percibida por el alumno

Nivel 1:
estudiante

COMPETE	Competitividad
EMOSUPS	Apoyo emocional de los padres percibido por el alumno
BELONG	Bienestar subjetivo: sentimiento de pertenencia a la escuela
DISCRIM	Ambiente escolar discriminatorio
WB176Q01HA	¿Cuándo fue la última vez que hiciste los deberes/estudiaste para el colegio?
MASTGOAL	Orientación a objetivos de dominio
GFOFAIL	Miedo general al fracaso
EMOSUPS	Apoyo emocional de los padres percibido por el alumno
EC154Q01IA	Clases de ampliación de Lengua
EC154Q02IA	Clases de ampliación de Matemáticas
EC154Q03IA	Clases de ampliación de Ciencias
EC154Q04HA	Clases de ampliación de Lengua Extranjera
EC154Q01IA	Clases de refuerzo de Lengua
EC154Q01IA	Clases de refuerzo de Matemáticas
EC154Q01IA	Clases de refuerzo de Ciencias
EC154Q01HA	Clases de refuerzo de Lengua Extranjera
EC154Q01IA	¿Actualmente asiste a clases adicionales para mejorar sus habilidades de estudio?
EC158Q01HA	El último día que fuiste a la escuela, ¿cuántas horas estudiaste por la mañana antes de ir a la escuela?
EC158Q02HA	El último día que fuiste a la escuela, ¿cuántos minutos estudiaste por la mañana antes de ir a la escuela?
EC159Q01HA	El último día que fuiste a la escuela, ¿cuántas horas estudiaste después de salir de la escuela?
EC159Q02HA	El último día que fuiste a la escuela, ¿cuántos minutos estudiaste después de salir de la escuela?

Nivel 1:
estudiante

