Check for updates

Human Behavior and Emerging Technologies

WILEY

RESEARCH ARTICLE OPEN ACCESS

Digital Habits and Time Use Patterns: Factors Influencing Reading Comprehension Skills in Spanish Adolescents Based on PISA 2022

Juan Ramón Rico-Juan[1] | Óscar Navarro-Martínez[2] | Beatriz Peña-Acuña[3]

[1]Department of Programming Languages and Informatics Systems, University of Alicante, Alicante, Valencian Community, Spain | [2]Department of Pedagogy, University of Castilla la Mancha, Ciudad Real, Community of Castilla la Mancha, Spain | [3]Department of Philology, University of Huelva, Huelva, Community of Andalusia, Spain

Correspondence: Beatriz Peña-Acuña (beatriz.pa@dfilo.uhu.es)

Received: 8 April 2025 | **Revised:** 16 March 2026 | **Accepted:** 27 March 2026

Academic Editor: Nikhat Kaura

Keywords: digital habits | explainable AI | grade repetition | home environment | machine learning | PISA 2022 | reading comprehension | sport practice

ABSTRACT

Adolescents' reading comprehension skills are shaped by various activities and contextual factors. Identifying these determinants is imperative for improving academic performance. This study aims to assess how time spent on different activities—sports, smartphone use, social media, and educational apps—affects reading comprehension in Spanish adolescents. The study is based on PISA 2022 microdata, including responses from over 30,000 Spanish adolescents. We examined factors such as grade repetition, time spent on sports, access to books at home, smartphone and social media usage, and engagement with educational apps. We applied several machine learning algorithms using cross-validation and the R^2 metric to evaluate model performance and analyze data. Additionally, explainable artificial intelligence techniques are used to interpret the results and highlight key predictors and their interactions. CatBoost yielded the best predictive performance, identifying key predictors of reading comprehension. The most influential factors included grade repetition (negative impact), time spent on sports (negative impact), access to books at home (positive impact), smartphone use (positive impact), and social media use (negative impact). Educational apps were associated with lower comprehension skills scores. Digital tools in technology lessons enhance reading skills when used regularly. The study highlights complex relationships between technology use, sports, and reading comprehension skills. These insights can inform educational policies and interventions targeting balanced digital use and tailored support for students struggling with academic challenges.

Highlights

- Grade repetition and participation in sports negatively impact reading comprehension skills.
- Home libraries enhance adolescent literacy skills.
- Smartphone usage supports reading comprehension skills, but educational apps show negative effects.
- The influence of parents' educational level on reading comprehension among adolescents is diminishing.
- Excessive use of social media reduces reading comprehension skills in adolescents.
- Digital learning tools in computer science education improve reading comprehension skills.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Copyright © 2026 Juan Ramón Rico-Juan et al. *Human Behavior and Emerging Technologies* published by John Wiley & Sons Ltd.

1 | Introduction

The Programme for International Student Assessment (PISA) study consists of assessing the competencies of students worldwide, and in its 2022 edition, approximately 700,000 students from 80 countries participated in the PISA 2022 [1], randomly selected from a total population of around 29 million students.

In general, the reading comprehension assessment in PISA evaluates the ability to understand, use, evaluate, reflect on, and engage with texts. While specific cognitive processes such as inference making or information integration are crucial [2, 3], this study focuses on the global score for this dimension among adolescent participants based on the PISA study 2022. This global indicator synthesizes students' performance across all underlying reading-related processes, providing a single, comparable measure of overall proficiency. This approach allows us to analyze how external determinants—specifically daily habits and the technological environment—impact overall performance, shifting the focus from isolated cognitive subskills to the broader context of the student's lifestyle. Moreover, understanding how these broader factors interact with reading performance is essential for interpreting cross-national trends.

The reading comprehension skills dimension in PISA 2022 continues to show great interest in the psychological–cognitive approach.

Spain's situation regarding reading comprehension skills (noted in the report under the label reading) is still improvable, according to the PISA 2022 report [1]. Despite being a developed country, Spain ranks in the average range in the second quartile for reading comprehension scores, along with 12 other countries. Spain has a reading average score of about 481, placing it 28th in this ranking of 81 countries. This study has the advantage of focusing on a single country, thereby contextualizing the cultural factor into one.

Building on this conceptual framework and the evolving nature of reading assessment across PISA cycles, the present study analyzes PISA 2022 data using advanced analytical techniques.

These analytical tools allow us to explore how students' daily habits, as captured in the PISA 2022 questionnaires, relate to their reading comprehension skills.

Using the PISA 2022 dataset, this article presents an in-depth analysis of how Spanish adolescents allocate their time and the potential impact on their reading comprehension skills. Adolescents' daily lives are increasingly defined by a complex interplay between digital engagement and traditional activities. The ubiquity of smartphones and social media raises questions about their potential interference with reading habits, while educational apps are often promoted as learning aids. Simultaneously, traditional factors such as the availability of a home library or participation in sports continue to compete for students' time. Furthermore, structural academic indicators, specifically grade repetition, remain critical in the Spanish educational context.

To address the gap in understanding the combined effect of these modern and traditional habits, we selected a specific set of predictors from the PISA questionnaires to use in a predictive regression model, taking Spain as a case: grade repetition, sports practice, book availability at home, and time spent on smartphones, educational apps, social networks, and digital content creation. Although prior research has examined some of these factors individually, no study has comprehensively investigated them together using ML techniques with a focus on explainability. This approach can provide a better understanding of the relationships between these factors, thereby helping to identify habits or circumstances that should be either encouraged or discouraged. Furthermore, Spain is analyzed as a country-level case within the comparative statistical framework. Accordingly, the present study seeks to address the following research questions:

- RQ1. What are the key technological and nontechnological factors that predict reading comprehension skills in Spanish adolescents?
- RQ2. How does the time spent on these specific activities positively or negatively influence reading performance?

2 | Literature Review

2.1 | Contextual Factors Influencing Reading Comprehension Skills

Reading comprehension is not an isolated cognitive skill but is deeply influenced by the environment in which adolescents grow and learn. While cognitive factors such as vocabulary and inference making are foundational [2], large-scale assessments like PISA highlight the critical role of contextual variables. These include the home environment, leisure habits, and the integration of technology in daily life. Understanding these determinants is essential, as reading literacy remains a primary predictor of future academic and professional success [4].

PISA consistently shows that student-level factors such as metacognition, enjoyment of reading, self-concept as a reader, and the language spoken at home are among the strongest predictors of reading comprehension (Firat & Conyucu [5–9]). At the family level, socioeconomic status, the availability of books and learning resources, parents' educational level, and shared reading habits play a central role in shaping literacy outcomes [7–9]. Within the school and classroom context, variables such as classroom climate, discipline, teacher support, and the amount of time devoted to reading activities are repeatedly associated with higher reading performance [6, 8, 10, 11]. Finally, technological factors, including ICT competence, moderate and purposeful Internet use, and appropriate family monitoring, also influence reading comprehension, highlighting the importance of balanced digital engagement [6, 9, 11].

Previous research with adolescents has shown that lifestyle-related variables—such as digital habits, leisure preferences, and time-use patterns—can be as influential as traditional sociodemographic factors in shaping reading performance [11, 12]. These studies highlight the need to consider broader contextual determinants beyond parent′s educational level or school type when analyzing reading comprehension in large-scale assessments.

2.2 | Home Environment and Leisure Habits

A substantial body of research has emphasized the importance of home literacy environments, including shared reading practices,

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

access to printed materials, and opportunities for reading for pleasure [13, 14]. Recent research with adolescents shows that reading on paper supports better comprehension of main ideas compared to reading on screens, suggesting that traditional print materials may offer cognitive advantages during complex reading tasks [15].

A growing body of evidence from PISA analyses consistently highlights the importance of home and leisure environments in shaping adolescents' reading comprehension. One of the strongest predictors is the availability of books at home, which reflects a rich literacy environment and is repeatedly associated with higher reading performance in international models [16–19]. Similarly, higher parents' educational level tends to correlate with better reading outcomes, as it often translates into greater academic support and more sustained exposure to literacy practices [11, 17, 19, 20]. Leisure reading—particularly the habitual reading of books and fiction—also shows a robust positive relationship with PISA reading scores, reinforcing the value of voluntary engagement with texts [21–24]. In contrast, heavy engagement in leisure digital activities, especially recreational screen use and gaming, is frequently linked to lower reading performance [11, 18, 25], suggesting that these habits may displace time and cognitive resources needed for deep reading. These findings align with broader research traditions that have examined specific components of the home environment, such as the availability of books and shared reading practices.

2.2.1 | Availability of Books at Home

The presence of books in the household is a robust proxy for a family's scholarly culture and has been consistently linked to higher literacy levels. Sikora et al. [26], in a large-scale study covering 31 societies, demonstrated that growing up in a home with a sizable library enhances adult literacy, numeracy, and technological problem-solving skills, independent of parents' education or occupation. This "scholarly culture" promotes a context in which reading is valued, directly impacting adolescents' comprehension skills. Studies such as Merga [13] emphasized that shared reading practices and access to diverse printed materials foster positive attitudes toward reading and richer comprehension.

2.2.2 | Impact of Sports Practice

The relationship between physical activity and academic performance is complex. Some research suggests that physical activity can improve executive functions, which support reading comprehension [27, 28]. Tiberi et al. [29] argued that cooperative sports foster cognitive skills beneficial for learning. However, the time displacement hypothesis suggests that excessive time spent on extracurricular activities, such as intensive sports training, may reduce the time available for academic pursuits and reading, potentially leading to lower performance [30]. Other research highlights the role of motor coordination tasks in supporting language and discourse development [31, 32], as well as the importance of socialization and external support programs in improving both self-esteem and reading-related skills [33, 34].

2.3 | Technological Habits and Screen Time

Technological habits have become central to adolescents' daily lives, shaping how they access information and engage with learning. While excessive or unstructured screen time can be detrimental, evidence shows that not all digital practices have the same impact. A systematic review by Adelantado-Renau et al. [35] reports that leisure screen use, including television, social media, and video games, is negatively associated with academic achievement, and intensive social media use may hinder deep reading [36]. In contrast, specific forms of active digital engagement—such as content creation requiring synthesis and evaluation—can enhance reading comprehension in certain contexts [37, 38]. Overall, the literature differentiates between passive digital consumption, often linked to poorer outcomes, and active, purposeful engagement, which may offer cognitive benefits.

Across PISA studies, research on total ICT use consistently shows that high levels of overall screen exposure tend to produce neutral or negative effects on academic outcomes [12, 39, 40]. Evidence on excessive recreational use indicates that leisure-oriented screen time, including social media, gaming, and video streaming, is associated with poorer performance in reading, mathematics, and science [18, 41, 42]. In contrast, studies examining ICT attitudes and autonomy report positive and stable associations, suggesting that students who display confidence, self-regulation, and purposeful engagement with digital tools achieve higher levels of academic performance [12, 39, 43]. Finally, research on classroom distraction highlights that the use of social networks during lessons is linked to reduced attention and lower achievement [40, 44].

These patterns are further reflected in specific digital behaviors, such as smartphone use and engagement with educational applications.

2.3.1 | Smartphone Usage and Educational Apps

The ubiquity of smartphones has transformed how adolescents access information. While moderate use can support digital reading skills, the nature of the usage is critical. Johnson [45] found that mobile phones could assist adolescents with limited language proficiency. However, the use of specific educational apps outside of formal schooling shows mixed results. While some interventions suggest benefits for vocabulary retention [46], others point to a lack of clear evidence regarding their effectiveness for higher level comprehension compared to traditional methods.

2.3.2 | Social Networks and Digital Content Creation

The time spent on social media and digital entertainment is increasingly concerning. A systematic review by Adelantado-Renau et al. [35] shows that leisure screen time—television, social media, and video games—is inversely related to academic performance, and high social media engagement may compete with the cognitive resources needed for deep reading [36]. Studies with Spanish adolescents indicate that online reading or research can support learning, whereas excessive social media use and gaming predict poorer school outcomes [12], and intensive social media use is linked to weaker reading skills [43]. Conversely, creating digital content requires synthesis and evaluation skills (Turner et al., 2019), and some tools for multimodal tasks may enhance reading comprehension in specific

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

contexts [38], underscoring the distinction between passive and active digital engagement.

2.4 | School Context: Grade Repetition

Grade repetition is one of the most debated policies in PISA reports: Some countries prohibit or restrict it, while others make extensive use of it. As a structural feature of education systems, repetition has significant implications for academic achievement, and Spain stands out for having rates well above the OECD average. Evidence from PISA consistently identifies repetition as a strong predictor of lower academic outcomes. In Spain, several studies using causal methods (instrumental variables, selection models, and propensity-score techniques) show that repeating a year substantially reduces achievement, with estimated effects between 1.5 and 1.7 standard deviations in PISA 2003–2018 [47]. The broader PISA debate similarly concludes that repetition, as currently implemented, does not improve performance and instead exacerbates inequalities [48] and psychosocial risks [49]. Consequently, studies recommend early intervention, individualized support, and pedagogical adjustments rather than relying on repetition as the standard response to low achievement.

Beyond PISA, Marsh et al. [50] and Jimerson [51] show that repetition rarely addresses learning gaps and is instead linked to lower achievement and socioemotional difficulties. Research in the field of reading similarly indicates that retained students tend to exhibit persistent difficulties, and studies such as Demagistri [52] and Sainio et al. [53] associate repetition with both reduced academic performance and heightened emotional distress during adolescence. Motivation and engagement also play a key role: students with stronger reading skills typically display higher academic motivation [54], while stressful life events and impairments in executive function contribute to poorer outcomes among repeaters [55].

2.5 | Use of Machine Learning (ML) Models With Explainability in the Context of Data Analysis

ML, a core branch of artificial intelligence (AI), revolves around algorithms designed to automatically extract underlying data patterns and forecast future observations (supervised learning). This methodology fundamentally relies on a two-phase workflow: initially fitting a model on a training dataset and subsequently leveraging that learned structure to predict outcomes on unseen records. The robustness of ML has driven its successful integration across highly diverse domains. For instance, it has updated predictive maintenance and quality control in the industrial sector [56–58], optimized diagnostics and clinical analytics in healthcare [59, 60], and significantly advanced the field of educational data mining [43, 61, 62]. In the specific context of international educational assessments, recent works like Álvarez-García et al. [63] have successfully applied unsupervised ML techniques (such as clustering and dimensionality reduction) to the PISA 2022 dataset to categorize student profiles independently of their academic grades.

Despite their outstanding predictive capabilities, a significant barrier to the widespread adoption of advanced ML models is their inherent opacity, commonly referred to as the "black box" problem [64]. The analytical ecosystem encompasses a wide spectrum of paradigms—ranging from decision trees and support vector machines to neural networks and neighborhood-based algorithms. Their primary advantage lies in their flexibility to seamlessly model complex, nonlinear interactions within multidimensional datasets, which traditionally outperform strict linear assumptions [65]. To maximize predictive performance, the standard methodological consensus dictates benchmarking a variety of these algorithms to pinpoint the optimal architecture for the specific analytical problem at hand.

To bridge the gap between high accuracy and lack of transparency, the burgeoning field of explainable artificial intelligence (XAI) provides essential frameworks for interpreting complex models [66]. Among the most robust methodologies are post hoc interpretability algorithms, which are applied after model training to decipher its internal logic (XAI-ML). A prime standard within this approach is the SHapley Additive exPlanations (SHAP) framework, conceived by Lundberg [67]. By anchoring its calculations in cooperative game theory, SHAP systematically quantifies the exact contribution of each predictive feature, enabling researchers to comprehensively map variable influence across global datasets, specific demographic subgroups, or even individual adolescent profiles.

3 | Methods

3.1 | Participants and Study Design

The PISA is a cross-sectional, large-scale assessment administered by the OECD. It targets 15-year-old students to evaluate their knowledge and skills in mathematics, science, and reading literacy. The study employs a two-stage stratified sampling design: first, schools are selected proportional to size, and second, students are randomly selected within schools [1]. This design ensures a representative sample of the target population.

For this study, we utilized the Spanish subset of the PISA 2022 dataset[1]. After cleaning and preprocessing (described in Section 3.2.3), the final sample consisted of 30,800 students. The dataset includes student background questionnaires, school questionnaires, and cognitive assessment scores. The study design is observational, aiming to identify associations and predictive patterns rather than establishing causality through intervention.

Table 1 provides a summary of the data, indicating that the distribution of sampled students is relatively balanced by gender across most regions. The regions with the largest sample sizes are the Basque Country, Madrid, Navarre, and Galicia, each with over 1800 students. Regarding reading test performance, the regions with the highest average scores are Asturias and Castile and Leon, both achieving average scores exceeding 500 points.

3.2 | Data Description, Target Variable, and Predictors

The target variable for this study is the students' performance in reading comprehension (reading PISA score). The predictors were selected from the student and school context questionnaires based on the dimensions defined in our research questions: home environment, technological access and habits, and school factors.

Table 2 details the variables used. Educational levels are classified according to the International Standard Classification of Education (ISCED). The reliability of the self-reported scales is documented in the PISA technical reports [1]. These predictors

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Number of participants and descriptive statistics (mean and standard deviation) of reading comprehension scores in the study sample, separated by region and sex.

Region	Reading			Sex					
				Female			Male		
	M	SD	N	M	SD	N	M	SD	N
Asturias	502.3	84.5	1724	513.7	80.7	841	491.4	86.6	883
Castile and Leon	502.0	82.0	1733	512.0	81.4	874	491.8	81.3	859
Madrid	499.3	85.9	2138	510.8	83.1	1052	488.2	87.2	1086
Cantabria	497.9	81.2	1677	505.4	79.0	800	491.1	82.5	877
Aragon	490.9	85.3	1420	506.9	81.6	669	476.6	86.0	751
Galicia	489.5	81.8	1824	502.9	75.8	903	476.4	85.3	921
La Rioja	488.3	88.6	1394	507.0	82.9	671	471.0	90.2	723
Comunidad Valenciana	486.7	84.9	1625	498.8	81.5	806	474.8	86.6	819
Navarre	479.5	86.1	1826	492.1	84.0	893	467.4	86.5	933
Balearic Islands	476.0	83.2	1576	485.9	81.9	785	466.2	83.4	791
Murcia	475.1	86.5	1669	486.6	82.6	824	463.9	88.7	845
Extremadura	472.9	83.7	1685	484.0	77.7	857	461.5	88.1	828
Basque Country	472.1	83.5	3244	487.5	78.2	1616	456.7	85.8	1628
Castile-La Mancha	470.9	80.9	1671	481.4	78.0	819	460.8	82.3	852
Canary Islands	466.2	84.9	1686	476.8	79.7	842	455.5	88.5	844
Andalusia	465.6	87.7	1707	476.3	83.1	872	454.4	90.9	835
Catalonia	464.5	93.5	1597	481.5	89.6	786	448.0	94.3	811
Ceuta and Melilla	421.3	87.4	604	430.0	84.2	329	410.9	90.1	275
Summary	480.9	86.2	30,800	492.9	82.6	15,239	469.0	88.1	15,561

were selected because they directly represent students' daily habits, home resources, and technological engagement, which align with our research questions.

3.3 | Data Analysis

For our analysis, we adopted an ML approach by testing and validating several algorithms. We selected a specific metric, commonly used in this type of study, to evaluate the results. This metric provided a robust framework for assessing the performance of the different algorithms under consideration.

3.3.1 | ML Approach

Recent methodologies in data analysis using ML models on PISA datasets have been demonstrated in studies such as those by Tan and Cutumisu [68], where a couple of ML algorithms are tested and compared, and in Miao et al. [69], which employs a single algorithm throughout the analysis. Similarly, our research follows the approach outlined by Rico-Juan et al. [11], aiming to improve the robustness of results by evaluating a range of ML algorithms that represent distinct learning paradigms proven to yield the best outcomes. This comprehensive approach allows us to select the most suitable algorithm for predicting outcomes, thereby enhancing both accuracy and interpretability of the results. The algorithms examined are categorized into linear models, decision and random trees, boosting, gradient boosting techniques, and neural networks.

To ensure a comprehensive evaluation, we benchmarked a diverse array of predictive architectures against a naïve Baseline (which systematically outputs the training set's mean reading score). The evaluated methods encompass several algorithmic families:

- Linear approaches: This includes standard linear regression fitted via least squares, alongside regularized variants designed to manage multicollinearity, such as Ridge [70], LASSO [71], and Automatic Relevance Determination (ARD [72]).
- Tree-based structures: Ranging from a single decision tree, which splits data based on optimal feature thresholds [73], to random forests, which build an ensemble of unpruned trees to significantly reduce variance [74].
- Boosting ensembles: We tested AdaBoost, which iteratively adjusts the weights of sequential linear regressors based on prior prediction errors [75]. Additionally, we evaluated state-of-the-art gradient boosting algorithms—specifically XGBoost [76], CatBoost [77], LightGBM [78], and Histogram GBR [79]. These methods sequentially build decision trees by minimizing differentiable loss functions, typically offering highly accurate predictions for tabular data.
- Neural Networks: Finally, we included a multilayer perceptron (MLP), representing a classic feedforward neural network with fully connected successive layers [80].

This diverse set of algorithms provides a comprehensive framework for identifying the most effective models, allowing for enhanced predictions and deeper insights into the analyzed data.

To evaluate the predictive performance of the employed ML models, we adopt an evaluation framework. Specifically, we

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Dimensions and variables used in this study (PISA Code, description [missing %], type, and values).

Dimension	PISA code	Description (missing %)	Type	Values
Home	ST007Q01JA	Father's highest level of schooling (7)	Cat	(ISCED 3.4, ISCED 3.3, ISCED 2, ISCED 1, ISCED < 1)
	ST005Q01JA	Mother's highest level of schooling (6)	Cat	(ISCED 3.4, ISCED 3.3, ISCED 2, ISCED 1, ISCED < 1)
	ST253Q01JA	Digital devices with screens count at home (3)	Cat	(No screens, One, Two, Three, Four, Five, 6–10, 10+)
	IC177Q06JA	Time spent: informational materials (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
	ST296Q04JA	Time spent on total homework (7)	Cat	(≤ 30 min, 30 min-1 hr, 1–2 h, 2–3 h, 3–4 h, 4+ hrs)
Technological	IC173Q04JA	Digital resource use in computer science lessons (8)	Cat	(Never, < 1/2 lessons, 1/2 lessons, > 1/2 lessons, Every lesson, No subject)
	IC171Q02JA	Time spent: out-of-school smartphone use (10)	Cat	(Never, 1–2 times/month, 1–2 times/week, 1 times/day, 1+ times/day, Not available)
	IC171Q04JA	Time spent: out-of-school internet access (11)	Cat	(Never, 1–2 times/month, 1–2 times/week, 1 times/day, 1+ times/day, Not available)
	IC171Q01JA	Time spent: out-of-school computer use (10)	Cat	(Never, 1–2/month, 1–2/week, Daily, Several/day, Not available)
	SC195Q02JA	Time spent: share digital content on social networks or platforms (11)	Cat	(Never, 1–2/yr, 3–5/yr, Monthly, 1+/month)
	IC171Q03JA	Time spent: out-of-school tablet use (10)	Cat	(Never, 1–2 times/month, 1–2 times/week, 1 times/day, 1+ times/day, Not available)
	IC171Q06JA	Time spent: out-of-school games use (10)	Cat	(Never, 1–2 times/month, 1–2 times/week, 1 times/day, 1+ times/day, Not available)
	IC177Q03JA	Time spent: browse internet for fun (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
	IC177Q01JA	Time spent: play video games (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
	IC177Q02JA	Time spent: browse social networks (11)	Cat	[No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day]
	IC171Q05JA	Time spent: out-of-school educational apps use (10)	Cat	(Never, 1–2 times/month, 1–2 times/week, 1 times/day, 1+ times/day, Not available)
	IC177Q05JA	Time spent: social networks (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
	IC177Q04JA	Time spent: look for practical information online (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
	IC177Q07JA	Time spent: create or edit digital content (11)	Cat	(No time, < 1 h/day, 1–3 h/day, 3–5 h/day, 5–7 h/day, 7+ hrs/day)
Skill perception	CURIOAGR	Curiosity (agreement) (WLE) (6)	Num	(−5: +5)
	PERSEVAGR	Perseverance (agreement) (WLE) (5)	Num	(−6: +6)
	ASSERAGR	Assertiveness (agreement) (WLE) (7)	Num	(−8.5: +8.5)
Geographic	REGION	Region (0)	Cat	(Andalusia, Aragon, Asturias, Balearic Islands, Canary Islands, Cantabria, Castile and Leon, Castile-La Mancha, Catalonia, Extremadura, Galicia, La Rioja, Madrid, Murcia, Navarre, Basque Country, Comunidad Valenciana, Ceuta & Melilla)
Scholar	SCHTYPE	School type (1)	Cat	(Public, Public-Private, Private)
	REPEAT	Grade repetition (3)	Cat	(Never, Repeated)
Others	ST004D01T	Gender (0)	Cat	(Female, Male)
	ST294Q05JA	Time spent: Exercise or practice a sport (4)	Cat	(0 days, 1 day, 2 days, 3 days, 4 days, 5+ days)
Reading	PVREAD	Reading PISA score (0)	Num	(0:1000)

Note: WLE stands for weighted likelihood estimates. Cat and Num values of column type correspond to categorical and numerical.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

implement the 10-fold cross-validation method, a widely accepted approach in statistical modeling that enhances the reliability of evaluation by partitioning the dataset into 10 disjoint subsets. Each subset is used once as a validation set, while the remaining subsets form the training set. This approach ensures that the models are evaluated on unseen data, providing a more accurate estimate of their predictive performance.

To compare the performance of different models, we employ the R-squared (R^2) metric, a commonly used measure in regression analysis [81]. R^2 represents the proportion of variance in the dependent variable that is predictable from the independent variables, providing a measure of how well the model explains the variability in the data [82]. The R^2 metric ranges from 0 to 1, where higher values indicate better model performance, with 1 representing a perfect fit. This intuitive measure of model goodness-of-fit enables straightforward comparisons between models, allowing us to identify the most effective models for predicting the target variable.

3.3.2 | Preprocessing Data

Many traditional ML algorithms require that all input data contain known values; otherwise, the algorithms cannot be applied. To resolve the issue of missing data, two main strategies can be employed: removing samples with missing values (which in our case would result in the loss of approximately 33% of the dataset) or applying imputation techniques. In our study, as shown in Table 2, the dataset contains missing values, and to preserve as much information as possible for model training, we chose to apply an imputation algorithm.

We selected an advanced multivariate imputation method [79, 83]. This method operates iteratively in a sequential updating process (round-robin), where each feature with missing values is predicted using the information from other observed variables. The imputation process involves constructing regression models for each feature, using the available data from other features to estimate and fill in missing values. This approach ensures that all observed data, as well as the relationships between features, are fully utilized during model training.

Regarding data standardization, numerical features in the PISA dataset are supplied as weighted likelihood estimates (WLE) [84], meaning that they do not require any additional scaling for our models. Categorical variables, however, were transformed into binary vectors via one-hot encoding. This procedure expands each categorical predictor into several distinct Boolean columns (one for each possible category level). In this format, a value of one indicates the presence of that specific trait for a given student, while the remaining related columns are set to 0. This vectorization is particularly necessary to facilitate the mathematical convergence of linear or distance-based algorithms, although tree-based models can inherently handle categorical splits regardless of this transformation.

By transforming the original data into these common representations, we enhance the learning capabilities of certain types of algorithms, particularly linear algorithms. For algorithms based on decision trees, for example, this transformation does not affect their learning.32

3.3.3 | ML and Explainability

High predictive accuracy in ML usually comes at the cost of algorithmic opacity. To unveil the internal decision-making process of our best-performing model, we implemented post hoc explainability techniques (see Figure 1).

Traditionally, variable importance is estimated by measuring performance drops after randomly shuffling a specific predictor (permutation method) [74]. However, we opted for a more mathematically robust framework based on Shapley values [85]. By calculating the average marginal contribution of a feature across all possible predictive combinations, SHAP guarantees key theoretical properties such as local accuracy and consistency. This theoretically sound metric translates complex algorithmic pathways into interpretable impacts, allowing us to accurately quantify how specific digital habits, sports practice, or home resources alter reading comprehension scores in adolescents.

Shapley values are particularly useful when understanding the overall impact of a variable is necessary, but explaining its influence on a group of samples or a specific sample is challenging. Lundberg and Lee [86, 87] have made significant progress in addressing these challenges. One valuable tool that follows this approach is SHAP [67], which enables explainability at the global, group, or individual level.

4 | Results

This section presents the findings derived from the ML analysis. First, we validate the selection of the CatBoost model. Second, we address RQ1 by identifying the most significant predictors of reading comprehension (Section 4.2). Finally, we answer RQ2 by analyzing the specific positive or negative influence of these time-use habits (Section 4.3) and statistically validating the profiles associated with high performance (Section 4.4).

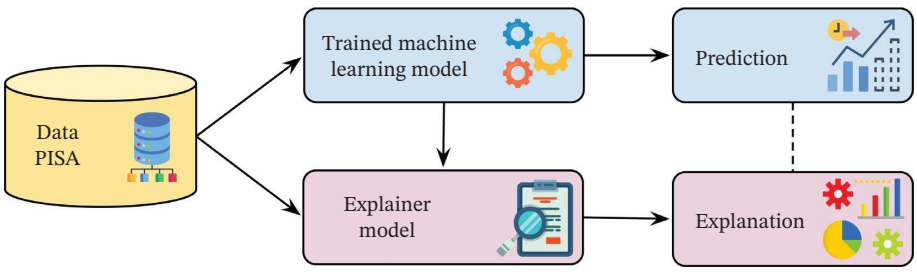


FIGURE 1 | General scheme of the post hoc explicability of machine learning models.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

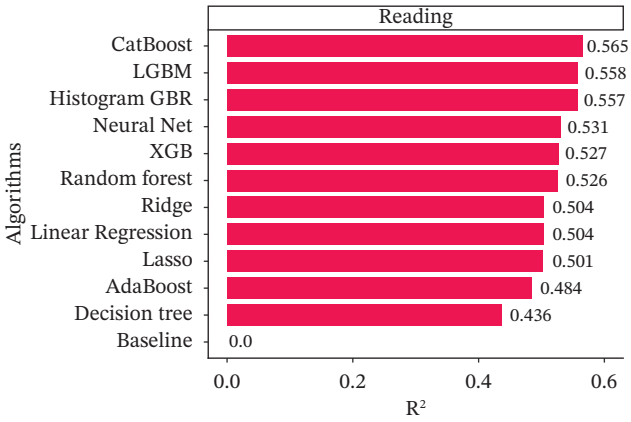


FIGURE 2 | Average results of 10-fold cross-validation using the R^2 metric. Higher results are better.

4.1 | Selection of the ML Model

To answer the research questions, we first need to select an ML model.

Figure 2 presents the average results of the cross-validation using the R^2 metric. Notably, the baseline model exhibits a value of zero, as this reference model does not account for any variance in the predictions by consistently predicting the average reading score of the training set. Furthermore, it is important to compare the results to ascertain whether the differences between the algorithms are statistically significant.

Figure 3 shows a comparison of the statistical significance between pairs of algorithms according to the R^2 metric. These comparisons were conducted using the Wilcoxon signed-rank test [88], employing the standard 95% confidence level [89] with Holm's adjustment [90]. This nonparametric statistical test does not assume a normal distribution of means. The results indicate that gradient-based tree and neural network models significantly outperform others (as denoted by the green bullets) as in Rico-Juan et al. [11]. Given the advantage of CatBoost over the others, we henceforth select it as the algorithm to perform explainability tests.

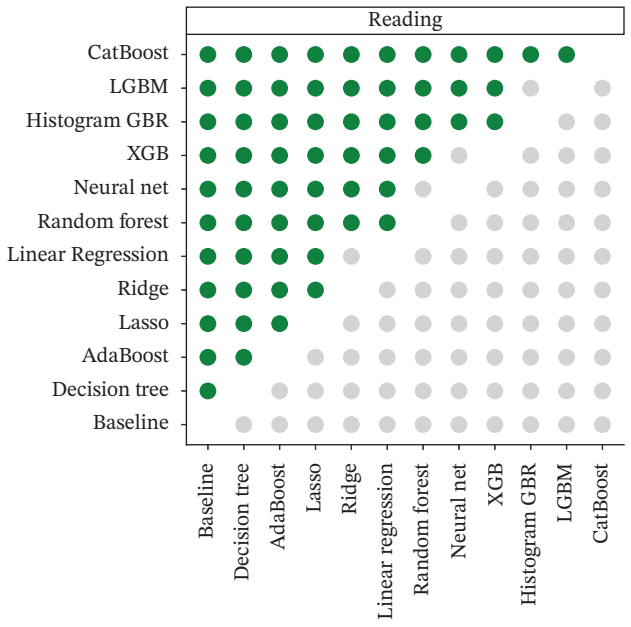


FIGURE 3 | Results of the Wilcoxon 10-fold cross-validation test according to the RMSE values. Green bullets indicate that the row algorithm is significantly better than the column algorithm at 95% confidence.

In addition, we conducted preliminary tests on some partitions using hyperparameter optimization algorithms [91] on models based on gradient boosting. Although this approach involves a high computational cost compared to learning with default parameters, the potential improvement in results can justify the parameter adjustment. However, in this case, the differences obtained were not greater than those achieved with the default parameters. Consequently, we decided not to perform the full tests with cross-validation.

4.2 | Identification of Key Technological and Nontechnological Factors (RQ1)

The significance of the variables is determined by their ability to explain the target variable, reading comprehension skills. The CatBoost algorithm is employed to quantify these relationships, as it efficiently incorporates categorical predictors without requiring extensive preprocessing and offers robust performance in high-dimensional analytical settings.

In the trained model, Shapley values are computed for each variable, with their mean absolute values used to quantify the relative importance of each predictor using a 10-fold cross-validation strategy. Figure 4 presents the predictors ranked according to their significance within the model. The subsequent analysis provides estimated impacts as determined by SHAP values, disaggregated by sample, with a focus on the most influential variables. It is relevant to contextualize these impacts within the range of theoretical PISA scores, which span from 0 to 1000. Empirically, in Spain, the highest average reading comprehension score is observed in Asturias (502.3 points), while the lowest average score is reported in Catalonia[2] (464.5 points), resulting in a notable disparity of 37.8 points between these regions (refer to Table 1). This performance gap serves as a benchmark for evaluating the estimated impacts of the variables presented in Figure 4 and subsequent figures. Additionally, it is important to consider that the difference of approximately 30 points in PISA scores is often interpreted as roughly equivalent to the academic progression of one school year [4].

The variables with the highest average impact are grade repetition with an absolute average impact of 13.8, and *Time spent: Exercise or practice a sport* with 13.38. This value represents slightly more than one-third of the 37.8-point difference observed between the top and bottom performing Spanish regions, and it also corresponds to about one-third of the difference in an academic course level.

We will statistically analyze the relationship between these two main variables by counting the number of adolescents in each category (Table 3). To do this, we will use Pearson's chi-squared test [92], yielding a result of $\chi^2\,(1, N = 30{,}800) = 655.24$, p-value $<$ 0.001. This allows us to reject the null hypothesis and conclude that these variables are not statistically independent, suggesting that there is a relationship between sports participation and grade repetition. Students who do not participate in sports (Sport: 0

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

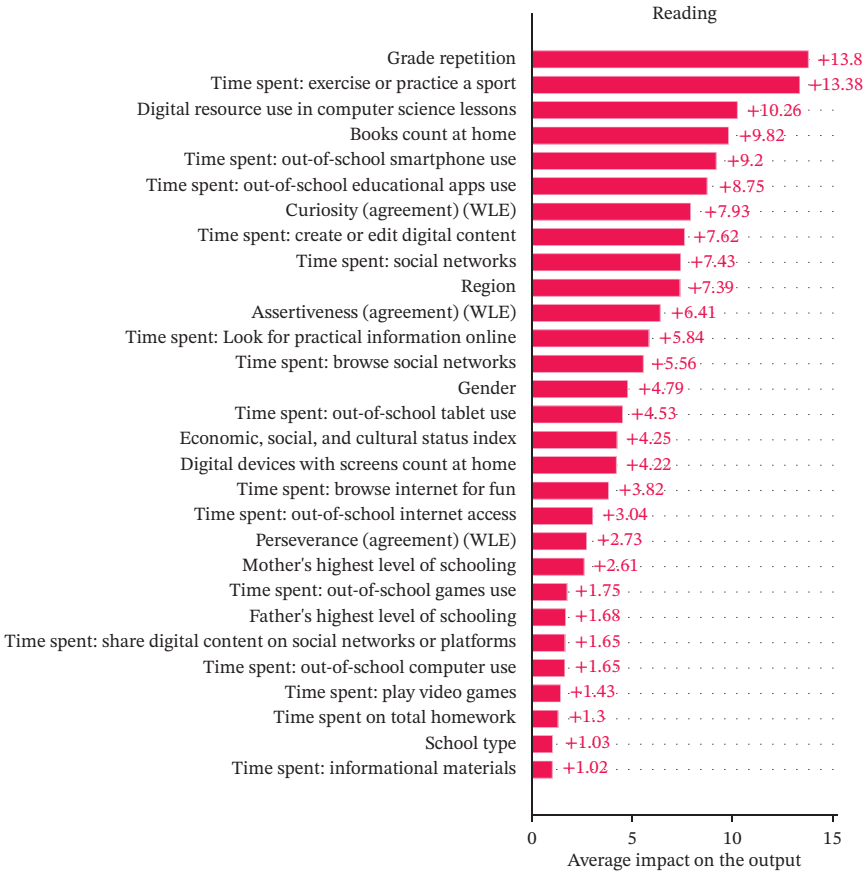


FIGURE 4 | Average importance in the absolute value of predictors and their impact on the outcome variable Reading.

days) appear more likely not to have repeated a grade, while those who do participate in sports (Sport: 0+ days) seem more likely to have repeated a grade. All of these results are related to RQ1.

It is worth noting that while psychological traits such as curiosity appear as relevant predictors (Figure 4), this study focuses specifically on modifiable habits and time allocation. Therefore, the subsequent detailed analysis will prioritize technological and daily activity factors over personality traits, as these offer more direct avenues for educational intervention.

4.3 | Analysis of the Influence of Time Allocation on Reading Skills (RQ2)

We will conduct a detailed analysis of the most relevant factors presented in the previous subsection. We focus on analyzing specific habits, such as how adolescents spend their time and how the home environment can positively or negatively influence this process. We will consider the frequency and type of activities to identify behavior patterns that either promote or hinder the development of these key academic competencies.

TABLE 3 | Number of Spanish adolescents with respect to grade repetition vs. time spent exercising or practicing a sport.

Time spent: Exercise or practice a sport	Grade repetition		
	Never	Repeat	Total
0 days	13,146	1950	15,096
0+ days/week	11,887	3817	15,704
Total	25,033	5767	30,800

4.3.1 | Impact of Time Spent in Sport Practice and Availability of Books at Home

Some variables show a positive relationship with the development of reading comprehension, particularly when their frequency of exposure increases (Figure 5). We observe that as the number of days per week that sports are practiced increases, it is related to obtaining lower results in reading skills. Adolescents who do not participate in sports score between 10 and 20 points higher. This score then decreases proportionally up to 3 days, where the trend stabilizes. Students who practice sports for three or more days score between 10 and 20 points lower.

Additionally, to estimate the causal effect of time spent in sports practice on reading scores, we employed double machine learning (DML) [93], an advanced statistical technique designed to estimate causal relationships more accurately. Unlike traditional methods, DML uses two stages to reduce bias and account for potential confounders. In the first stage, it estimates the relationships between the treatment variable (time spent in sports practice), covariates (other variables, Table 2), and the target (reading) using flexible ML models. In the second stage, it corrects for any bias in these estimates to isolate the true causal effect. The results from this approach[3] revealed a negative impact on reading scores as the number of days spent on sports practice increased. Specifically, the average treatment effects were −12.9 for 1 day per week, −17.7 for 2 days, −22.6 for 3 days, −22.8 for 4 days, and −25.0 for 5 or more days, all relative to 0 days of

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

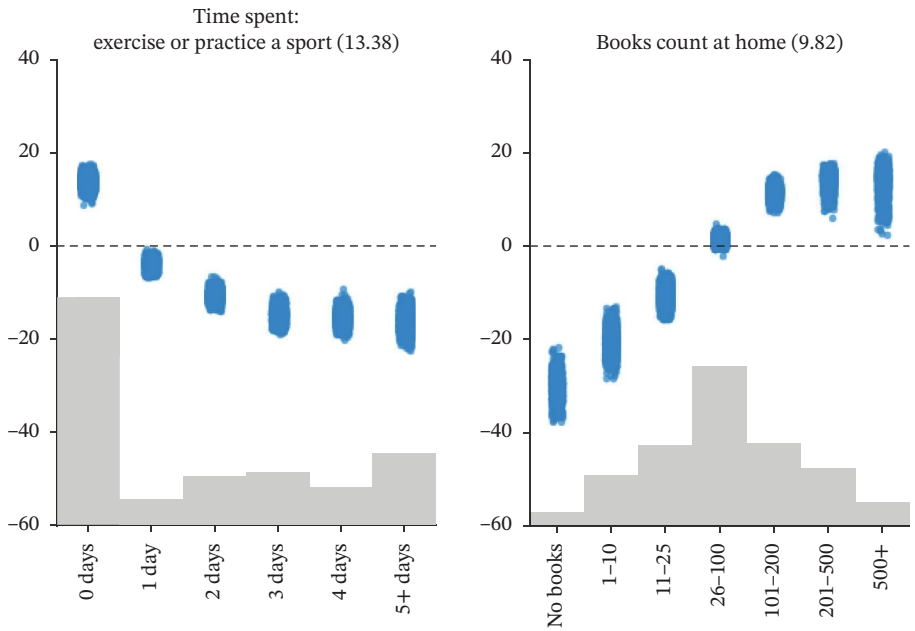


FIGURE 5 | The effects of the model on the frequency of practice sports and books at home variables. Each subplot shows the name of the variable and the average impact in absolute value in parentheses.

sports practice. This trend clearly indicates that, in similar circumstances, the more days a student spends on sports practice, the greater the negative impact on their reading scores. By using DML, we can be more confident that this relationship is causal and not merely correlational.

We also notice that adolescents who have a greater number of books at home tend to achieve better scores in this skill. Those without books at home show clearly inferior performance, with scores between 20 and 40 points below average. As the number of books increases, the scores also rise proportionally, reaching a notable increase up to a threshold of between 100 and 200 books. Beyond this point, the results tend to stabilize, achieving additional improvements of between 15 and 20 points in most cases. Both variables show opposing impacts on reading comprehension outcomes. On the one hand, the time spent engaging in sports negatively affects reading skills, with an average impact value of 13.28, while the increasing number of books in the home positively influences reading skills, with an average impact value of 9.89.

4.3.2 | Time Spent on Smartphone Usage and Use of Educational Apps Outside of School

The use of technology outside the school environment may also be relevant (Figure 6). The use of smartphones outside of school appears to correlate positively with reading comprehension. In

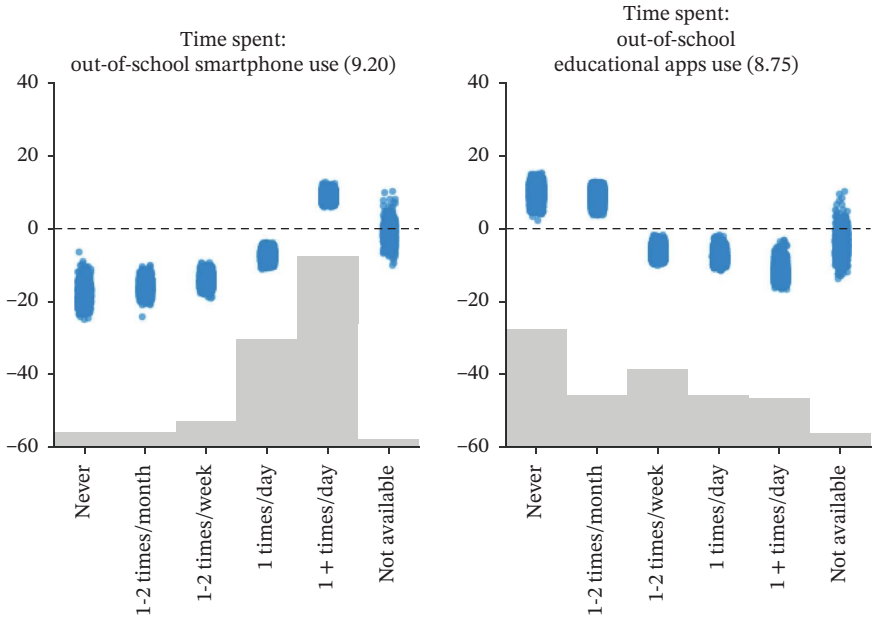


FIGURE 6 | The effects of the model on time spent out of school using a smartphone or educational apps. Each subplot shows the name of the variable and the average impact in absolute value in parentheses.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

cases where these devices are not used or are used with a frequency of less than once or twice a week, scores are observed to be between 10 and 25 points below average. However, when usage increases to once a day, the results improve, and more frequent use, that is, more than once a day, is associated with a slight additional gain of between 5 and 10 points.

On the other hand, the use of educational apps has the opposite relationship. The best results are obtained when these apps are not used, or when they are used only 1 or 2 times a month (between 5 and 15 points higher). If the frequency of use of these applications is greater, it results in a score of between 0 and 10 points lower, or even up to almost 20 points lower when used several times a day. The extracurricular use of smartphones appears to have a positive effect, with an average value of 9.21. In contrast, the use of educational applications seems to negatively impact reading skills, with an average value of 8.67.

4.3.3 | Time Spent Using Social Networks and Time Spent Creating Digital Content

Finally, in Figure 7, the impact of increased frequency of use on the same day of social media or digital content editing on the development of reading comprehension skills can be observed. In the case of social media, using it for less than one hour a day implies a greater reading comprehension, up to almost 20 points more if not used at all, and up to 10 points more if used for less than one hour. In contrast, when used for between 1 and 7 h, it reflects a score between 0 and 15 points lower. Furthermore, when used for more than 7 h a week, it can lead to a difference of up to 20 points lower.

When students spend time creating digital content, the results are very similar to the previous case. The trend is similar, although with some different nuances. We clearly obtain a higher score when this task is not performed at all, between 5 and 15 points more. There is practically no effect when used for less than one hour a day, and it starts to have relevance when digital content editing is done for between one and three hours, resulting in between 5 and 20 points lower in reading comprehension. This negative trend continues until these activities are performed for more than 7 h a day, which can lead to a reduction of up to 30 points. In both cases, a similar trend is observed: the greater the usage, the lower the reading comprehension. Additionally, the average values are practically identical, 7.43 and 7.62, respectively, for social media use and digital content editing.

4.3.4 | Analysis in the School Context: Grade Repetition and the Use of Digital Resources in Computer Science Lessons

In the previous sections, we have focused on analyzing variables that evaluate the habits of adolescents, particularly in relation to the use of digital resources. However, other factors inherent to the school environment also have a significant influence on the development of reading comprehension skills.

One of these factors is grade repetition, which shows a marked difference in the scores obtained in reading comprehension. According to the data presented in Supporting Information (see Supporting Information 1), students who have repeated at least one grade exhibit a decrease of between 30 and 50 points in their scores. In contrast, those who have not repeated achieve between 5 and 15 points higher. It is noteworthy that this variable has the highest average value among those studied, with an average impact in absolute value of 13.65 points.

Additionally, the use of digital resources in technological subjects shows a relevant impact. Students who use these resources in all academic sessions achieve between 0 and 10 additional points in their reading comprehension assessments. On the other hand, those who use them in less than half of the sessions experience a decrease in their scores, with losses ranging from 5 to 20 points. This variable also presents a high average, with a mean difference of 10.18 points.

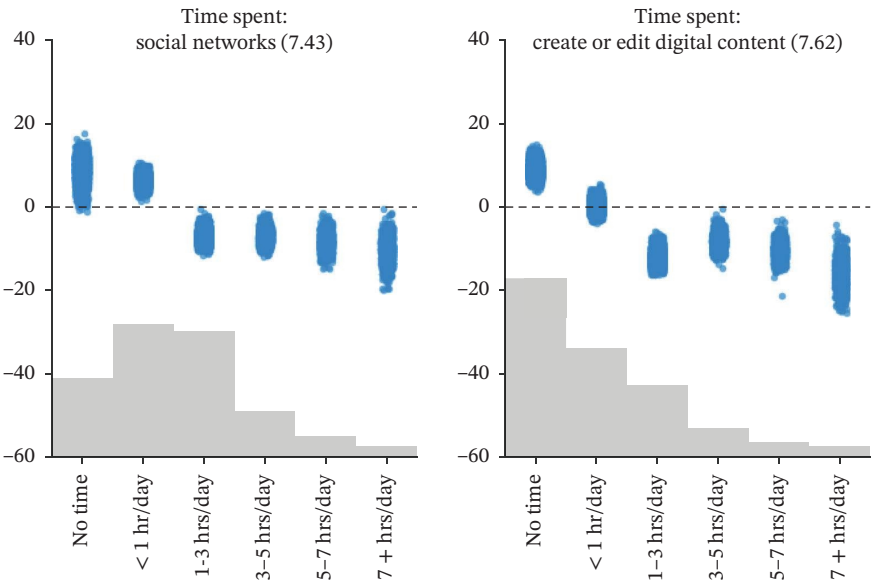


FIGURE 7 | The effects of the model on the time spent in social networks and working with digital content. Each subplot shows the name of the variable and the average impact in absolute value in parentheses.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Profiles showing positive impact on reading performance, analyzed using the Mann–Whitney U test.

Variable	Satisfy the criterion			Do not satisfy the criterion	U	p-value
	Description	N (%)	M	M		
Grade repetition	Never	81.3	499.1	401.6	117811693	< 0.001
Time spent: Exercise or practice a sport	0 days	49.0	510.0	452.9	164104501.5	< 0.001
Digital resource use in computer science lessons	No subject	24.8	517.6	468.8	117949126.5	< 0.001
Books count at home	+100 books	35.0	515.1	462.5	146313240.5	< 0.001
Time spent: out-of-school smartphone use	1+ times/day	52.5	510.0	448.7	167022852.5	< 0.001
Time spent: out-of-school educational apps use	< 1–2 times/month	14.3	495.0	478.5	64838753	< 0.001
Curiosity (agreement) (WLE)	+0	47.9	498.9	464.3	145281040	< 0.001
Time spent: social networks	< 1 h/day	50.8	507.3	453.6	161646575	< 0.001
Time spent: create or edit digital content	< 1 h/day	69.2	501.8	434.0	145958169	< 0.001
	All criteria	0.3	582.8	480.6	2312595	< 0.001

Note: N (%) indicates the percentage of students meeting the condition.

4.4 | Statistical Validation of High-Performance Profiles

The conditions associated with higher performance are based on the variables analyzed. Table 4 presents a set of statistically validated profiles derived from positive impact scenarios, as illustrated in the detailed impact graphs (Figures 5, 6, 7). These profiles were evaluated using the nonparametric Mann–Whitney U test [94], which compares reading test scores between two groups. The first group consists of students who meet the specified condition, while the second includes those who do not. As shown in the table, students who satisfy the conditions exhibit significantly higher performance levels, with a confidence level exceeding 99%. In the particular case, when all criteria are combined (bottom row, representing 0.3% of the total), the mean score reaches 582.8, surpassing the values for any individual criterion and exceeding the average performance observed in Spain.

5 | Discussion and Conclusion

5.1 | Interpretation of Key Factors Related to Reading Comprehension

Our analysis using XAI reveals a shift in the determinants of reading comprehension compared to previous PISA cycles such as Rico-Juan et al. [11] based on PISA 2018. While parents' educational level remains a factor, our model highlights that daily habits and time allocation have become dominant predictors for Spanish adolescents. This suggests that the *what* and *how much* of daily activities—technological or otherwise—play a more decisive role than purely socioeconomic background in the current context.

5.1.1 | Impact of Sports Practice and Home Library

Our results identify a robust negative relationship between the frequency of sports practice and reading comprehension scores. This finding, reinforced by our causal inference analysis (DML), contrasts with studies that emphasize the cognitive benefits of physical activity [28, 29]. We interpret this discrepancy through the time displacement hypothesis: For adolescents, the significant time commitment required for frequent sports training may directly compete with the time available for academic reading and study. While sports may enhance executive functions, this general cognitive benefit does not appear to compensate for the reduction in specific reading time in terms of PISA test performance.

Conversely, the number of books at home positively predicts reading skills, aligning perfectly with established literature [26]. The presence of a home library serves as a proxy for "scholarly culture," fostering an environment where reading is valued and encouraged, thus validating that traditional literacy resources remain vital even in a digital age. This interpretation aligns with PISA-based research, showing that activities which reduce exposure to sustained reading practices tend to be associated with lower reading outcomes, particularly when they displace time needed for deep, text-based engagement [18, 25, 43].

This result is fully consistent with recent PISA analyses, demonstrating that the availability of books at home remains one of the strongest predictors of reading performance across countries, even after controlling for socioeconomic status and parents'

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

educational level [16, 17, 19]. Moreover, our findings reinforce the idea that home literacy environments not only provide material resources but also shape adolescents' leisure habits, encouraging voluntary reading—an activity repeatedly shown to enhance PISA reading scores [21–24].

5.1.2 | Smartphone Usage and Educational Apps

Regarding mobile technology, we found a nuanced result: moderate smartphone usage shows a positive association with reading skills, whereas the use of educational apps outside of school is negatively associated. The positive link with general smartphone use may be attributed to the text-rich nature of many mobile interactions (messaging, information seeking), which can support reading fluency [45]. This interpretation is compatible with PISA-based findings, showing that moderate, purposeful ICT engagement—particularly when it involves reading or information-seeking behaviors—is associated with better academic outcomes, in contrast to the detrimental effects of excessive recreational screen exposure [12, 39].

However, the negative impact of educational apps is striking. This finding contradicts the assumption that "educational" labels guarantee learning gains [95]. A possible explanation is reverse causality: Students with lower reading skills may be directed by parents to use these apps as a remedial measure. Alternatively, many apps may prioritize gamification and visual engagement over deep textual processing, failing to develop the complex comprehension skills measured by PISA. This aligns with PISA research indicating that not all forms of digital engagement are academically beneficial: While ICT confidence and autonomy predict higher performance, the use of digital tools that do not require sustained reading or self-regulated learning tends to show neutral or negative associations [39, 43]. Moreover, the negative association mirrors findings on leisure-oriented screen time, where activities that emphasize rapid visual stimulation over text comprehension—such as gaming or video streaming—are linked to poorer reading outcomes [18, 41]. This highlights the need to scrutinize the pedagogical quality of such apps [96].

5.1.3 | Social Networks and Digital Content Creation

Our data indicate a clear decline in reading skills as time spent on social networks and digital content creation increases. This aligns with the view that excessive screen time for leisure fragments attention and reduces the capacity for deep, sustained reading [12, 36]. These results are consistent with PISA evidence, showing that high levels of recreational ICT use—particularly social media and video-based platforms—are associated with lower performance in reading, mathematics, and science [11, 18]. Furthermore, research on classroom distraction reinforces this pattern: The use of social networks during lessons has been linked to reduced attention and lower achievement [40, 44], suggesting that the negative effects extend beyond leisure contexts into academic environments.

While some authors argue that content creation fosters synthesis skills [38], our results suggest that in the context of Spanish adolescents, these activities largely act as distractors or time displacement factors rather than skill-builders. This interpretation is compatible with PISA findings, indicating that when digital activities compete with time for reading or studying, the net effect tends to be negative, regardless of their creative or expressive potential [39]. The critical threshold identified suggests that usage beyond minimal levels becomes detrimental.

5.1.4 | School Context: Grade Repetition and Digital Resources

Grade repetition emerged as the strongest negative predictor in our model. This confirms that repeating a year is not an effective remedial strategy for reading proficiency; instead, it is associated with a persistent gap in performance [50]. This structural failure in the system highlights the need for early intervention rather than retention. Our findings are fully aligned with PISA-based evidence, showing that repetition is one of the most powerful negative predictors of academic achievement across countries, and particularly in Spain, where causal studies estimate reductions of between 1.5 and 1.7 standard deviations in PISA reading performance [47]. Moreover, the broader PISA literature indicates that repetition not only fails to improve learning but also exacerbates educational inequalities [48] and increases psychosocial risks such as disengagement and reduced self-esteem [49]. Thus, our results reinforce the consensus that retention, as currently implemented, is an inefficient and inequitable policy, and that system-level responses should prioritize early intervention, individualized support, and pedagogical adjustments rather than grade repetition.

In contrast, the regular use of digital resources within computer science lessons is positively associated with reading comprehension. Unlike unstructured leisure use (social media), this guided, pedagogical use of technology appears to support learning. This supports the findings of Beek et al. [97], suggesting that technology is beneficial when integrated into a structured learning environment with clear educational goals. This pattern is consistent with PISA research, showing that structured, academically oriented ICT use—particularly when supported by teacher guidance—tends to have positive or neutral effects on performance, in contrast to the negative associations observed for recreational screen time [12, 39]. Furthermore, studies on ICT autonomy and self-regulated digital learning indicate that students who use technology purposefully and with clear instructional scaffolding achieve higher academic outcomes [43], which may help explain the positive association observed in our model.

The comparison between PISA 2018 and PISA 2022 reveals a shift in the relative weight of several contextual factors associated with reading performance. In PISA 2018, variables such as parents' educational level, gender, region, time spent on social media, and school type showed notable influence, with parental education emerging among the most impactful predictors. By contrast, PISA 2022 highlights a different configuration of relevant factors, with teacher support, the use of remote devices and digital platforms, socioeconomic status, and the language spoken at home gaining prominence. These changes suggest that while some traditional determinants of reading performance have diminished in explanatory power, others linked to technological engagement and learning environments have become more salient. This evolution reflects broader transformations in adolescents' educational and digital contexts and underscores the need to interpret reading outcomes within a changing landscape of social, technological, and instructional influences.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.2 | Contributions of the Study

This study offers several relevant contributions to the analysis of reading performance in the context of PISA. First, it integrates multiple dimensions—family environment, reading habits, technology use, sports practice, and grade repetition—into a single causal model, an approach that is uncommon in PISA-based research. This broad perspective provides updated empirical evidence for Spain, a country with distinctive characteristics such as high repetition rates and strong digital penetration.

The findings confirm the importance of the family environment, showing that the availability of books at home remains a strong predictor of reading performance and reinforcing the role of cultural capital and voluntary reading in the development of reading competence. The study also provides novel evidence regarding sports practice, identifying a negative effect associated with training frequency and contributing to the ongoing debate on its relationship with academic performance through the lens of the time displacement hypothesis.

With respect to digital habits, the study distinguishes between technological uses with clearly differentiated effects: While moderate smartphone use is positively associated with reading, the use of educational applications and the time devoted to social networks or digital content creation show negative associations. This differentiation between pedagogical and recreational uses of technology brings clarity to an area where findings are often heterogeneous and offers useful insights for informing educational policies that promote more effective and pedagogically grounded uses of digital resources.

Finally, in the school context, the study reaffirms that grade repetition is the strongest negative predictor of reading performance, in line with causal evidence available for Spain and with international PISA analyses. It also shows that the structured use of digital resources in the classroom, particularly in guided contexts such as computer science lessons, can have positive effects on reading comprehension, reinforcing the idea that educational technology is effective when integrated into well-designed pedagogical environments.

5.3 | Theoretical Contributions and Educational Implications

Building on these empirical findings, this study contributes to the field by quantifying the trade-offs between competing daily habits. In doing so, it establishes a conceptual link between the empirical patterns identified and broader theoretical debates, challenging the unqualified promotion of certain extracurricular or digital activities—such as intensive sports training or unguided use of educational applications—by demonstrating their potential opportunity costs for literacy development. These insights provide a clearer framework for interpreting how students' time allocation shapes academic outcomes, underscoring the need for educational policies that balance students' daily activities and promote structured, pedagogically meaningful engagement with both physical and digital practices.

Practical implications for educators and families include the following:

- Balanced Schedules: Families should be aware that while sports are healthy, high-frequency training without compensatory study time may hinder literacy development.
- Critical Tech Use: Educators should distinguish between "screen time" types. While guided digital use in class is beneficial, unstructured use (social media) and unverified "educational" apps require strict moderation.
- Early Intervention: The strong negative indicator of grade repetition calls for preventive support systems for struggling readers before retention becomes necessary.

The study provides empirical evidence that can inform public policy across several key areas. First, it highlights the need to strengthen policies that support families' cultural capital by promoting early access to reading materials and encouraging voluntary reading. The findings also point to the potential opportunity costs associated with certain extracurricular and digital activities, suggesting that educational policies should help balance students' use of time and promote guided, pedagogically meaningful technological practices. In the school context, the results support a review of grade repetition policy, given its strong negative association with reading performance, and indicate the value of replacing it with early interventions and individualized support. Finally, the study endorses the structured integration of digital resources in the classroom, particularly in guided contexts, as an effective means of improving reading comprehension.

5.4 | Limitations and Future Directions

While robust, this study is limited by the cross-sectional nature of PISA data (though mitigated by DML for sports) and the reliance on self-reported questionnaires. We could not assess the quality of the time spent (e.g., what books are read or which apps are used), only the frequency.

Future research should employ longitudinal designs to track how changes in these habits over time directly affect reading trajectories. Additionally, while this study identified key predictors using ML regression, future work could employ structural equation modeling (SEM) to map the causal interrelationships between these variables, effectively building the theoretical framework that this predictive analysis points toward.

Author Contributions

Juan Ramón Rico-Juan has participated in data curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; validation; visualization; roles/writing–original draft; and writing–review and editing. Óscar Navarro-Martínez has participated in conceptualization; formal analysis; investigation; methodology; supervision; roles/writing–original draft; and writing–review and editing. Beatriz Peña-Acuña has participated in conceptualization; formal analysis; investigation; methodology; supervision; roles/writing–original draft; and writing–review and editing in English.

Acknowledgments

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: During the preparation of this work, the authors used CHATGPT for PROOFREADING. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the published article.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Funding

This work was supported by the Spanish Ministry MCIN/AEI/(Ref. PID2024-161558OB-I00).

Disclosure

All authors approve the final version of the manuscript. Funding sources had no involvement.

Ethics Statement

This study relied exclusively on secondary open access data from the OECD's PISA database. The data are fully anonymized and publicly available, and therefore, no ethics committee approval or informed consent was required.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data used for this study are available at https://rua.ua.es/dspace/handle/10045/146991 (Institutional Repository of the University of Alicante) with public and open access. Specifically, the data relating to the *student questionnaire* and the *school questionnaire* in Spain available at https://www.oecd.org/en/data/datasets/pisa-2022-database.html have been merged.

Endnotes

[1]The microdata from the PISA 2022 report are publicly accessible at https://www.oecd.org/en/data/datasets/pisa-2022-database.html. We have used Spain as the country for our analysis.

[2]Ceuta and Melilla, with a score of 421.3, have been excluded due to their exceptional situation and continuing problems with respect to education.

[3]The EconML (v0.15.1), Catboost (v1.2.7), and scikit-learn (v1.5.2) libraries are used with the *LinearDML* estimator for double machine learning. The *CatBoostRegressor* is employed as the model for predicting the outcome variable, while *LogisticRegression* is used to model the multiclass treatment variable. The training process involves cross-validation with 5-fold to ensure robust performance.

References

1. OECD, "PISA 2022," (2023), https://www.oecd.org/en/about/programmes/pisa.html.

2. A. Elleman and E. Oslund, "Reading Comprehension Research: Implications for Practice and Policy," *Policy Insights from the Behavioral and Brain Sciences* 6, no. 1 (2019): 11–3, https://doi.org/10.1177/2372732218816339.

3. L. Salmerón, A. García, and E. Vidal-Abarca, "The Development of Adolescents' Comprehension-based Internet Reading Activities," *Learning and Individual Differences* 61 (2018): 31–39, https://doi.org/10.1016/j.lindif.2017.11.006.

4. OECD, *PISA 2018 Results. what Students Know and Can Do* (OECD Publishing, 2019), https://doi.org/10.1787/5f07c754-en.

5. T. Fırat and İ. Koyuncu, "The Influence of Student-Level Factors on Reading Literacy: A Comprehensive Study," *Journal of Theoretical Educational Science* 16, no. 4 (2023): 843–867, https://doi.org/10.30831/akukeg.1299077.

6. V. Özdemir, "Investigating Factors that Affect Reading Literacy Skills in PISA Turkey Sample," *Acta Educationis Generalis* 14, no. 1 (2024): 20–39, https://doi.org/10.2478/atd-2024-0002.

7. L. A. López-Agudo, S. M. González-Betancor, and O. D. Marcenaro-Gutiérrez, "Language at Home and Academic Performance: The Case of Spain," *Economic Analysis and Policy* 69 (2021): 16–33, https://accedacris.ulpgc.es/handle/10553/76068, https://doi.org/10.1016/j.eap.2020.11.003.

8. L. Rojas-Torres, G. Ordóñez, and K. Calvo, "Teacher and Student Practices Associated with Performance in the PISA Reading Literacy Evaluation," *Frontiers in Education* 6 (2021): 658973, https://doi.org/10.3389/feduc.2021.658973.

9. V. Vazquez-Lopez and E. Huerta-Manzanilla, "Factors Related with Underperformance in Reading Proficiency, the Case of the Programme for International Student Assessment 2018," *European Journal of Investigation in Health Psychology and Education* 11, no. 3 (2021): 813–828, https://doi.org/10.3390/ejihpe11030059.

10. İ. Koyuncu and T. Fırat, "Investigating Reading Literacy in PISA 2018 Assessment," *Lnternational Electronic Journal of Elementary Education* 13, no. 2 (2021): 263–275, https://doi.org/10.26822/iejee.2021.189.

11. J. R. Rico-Juan, B. Peña-Acuña, and O. Navarro-Martinez, "Holistic Exploration of Reading Comprehension Skills, Technology and Socioeconomic Factors in Spanish Teenagers," *Heliyon* 10, no. 12 (2024): e32637, https://www.cell.com/heliyon/fulltext/S2405-8440(24)08668-7, https://doi.org/10.1016/j.heliyon.2024.e32637.

12. D. Gómez-Baya, A. Grasmeijer, E. López-Bermúdez, M. Matos, and R. Mendoza, "Positive Youth Development and Internet Use in a Sample of Spanish Adolescents," *Frontiers in Pediatrics* 10 (2022): 842928, https://doi.org/10.3389/fped.2022.842928.

13. M. K. Merga, "Access to Books in the Home and Adolescent Engagement in Recreational Book Reading: Considerations for Secondary School Educators," *English in Education* 49, no. 3 (2015): 197–214, https://doi.org/10.1111/eie.12071.

14. K. Wilkinson, V. Andries, D. Howarth, J. Bonsall, S. Sabeti, and S. McGeown, "Reading During Adolescence: Why Adolescents Choose (Or do Not Choose) Books," *Journal of Adolescent & Adult Literacy* 64, no. 2 (2020): 157–166, https://doi.org/10.1002/jaal.1065.

15. A. Ronconi, V. Veronesi, L. Mason, et al., "Effects of Reading Medium on the Processing, Comprehension, and Calibration of Adolescent Readers," *Computers & Education* 185 (2022): 104520, https://doi.org/10.1016/j.compedu.2022.104520.

16. Ö. Güre, H. Şevgin, and M. Kayri, "Reviewing the Factors Affecting PISA Reading Skills by Using Random Forest and MARS Methods," *International Journal of Contemporary Educational Research* 10, no. 1 (2023): 181–196, https://doi.org/10.33200/ijcer.1192590.

17. E. Tekin and Y. Pınar, "Preschool Attendance, School Starting Age and Cultural-Financial Resources as Predictors of Adolescent's Reading Comprehension: Evidence from PISA 2015 and 2018," *Sage Open* 13, no. 4 (2023): https://doi.org/10.1177/21582440231207690.

18. E. Vázquez-Cano, J. Gómez-Galán, A. Infante-Moro, and E. López-Meneses, "Incidence of a Non-sustainability Use of Technology on Students' Reading Performance in Pisa," *Sustainability* 12, no. 2 (2020): 749, https://doi.org/10.3390/su12020749.

19. L. Paakkari, J. Ruotsalainen, H. Lahti, et al., "The Role of the Home in Children's Critical Reading Skills Development," *Humanities and Social Sciences Communications* 11 (2024): 1–10, https://doi.org/10.1057/s41599-024-02843-7.

20. C. Özcan and B. Gücüm, "Reading Comprehension Skills in Science: The Effect of Model-based Approach Supported by Parent Education Level, Home Environment and Reading Texts," *Reading Psychology* 46, no. 6 (2025): 583–605, https://doi.org/10.1080/02702711.2025.2482621.

21. E. Bergen, K. Vasalampi, and M. Torppa, "How Are Practice and Performance Related? Development of Reading from Age 5 to 15," *Reading Research Quarterly* 56, no. 3 (2020): 415–434, https://doi.org/10.1002/rrq.309.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

22. J. Jerrim and G. Moss, "The Link Between Fiction and Teenagers' Reading Skills: International Evidence from the OECD PISA Study," *British Educational Research Journal* 45, no. 1 (2018): 181–200, https://doi.org/10.1002/berj.3498.

23. M. Torppa, P. Niemi, K. Vasalampi, M. Lerkkanen, A. Tolvanen, and A. Poikkeus, "Leisure Reading (But Not Any Kind) and Reading Comprehension Support Each Other-A Longitudinal Study Across Grades 1 and 9," *Child Development* 91, no. 3 (2020): 876–900, https://doi.org/10.1111/cdev.13241.

24. E. Ulvinen, M. Psyridou, M. Lerkkanen, A. Poikkeus, M. Siekkinen, and M. Torppa, "Developmental Leisure Reading Profiles and Their Association with Reading Skills Across Grades 1–9," *Learning and Individual Differences* 109 (2024): 102387, https://doi.org/10.1016/j.lindif.2023.102387.

25. J. Gubbels, N. Swart, and M. Groen, "Everything in Moderation: ICT and Reading Performance of Dutch 15-year-olds," *Large-scale Assessments in Education* 8, no. 1 (2020): 1–17, https://doi.org/10.1186/s40536-020-0079-0.

26. J. Sikora, M. D. Evans, and J. Kelley, "Scholarly Culture: How Books in Adolescence Enhance Adult Literacy, Numeracy and Technology Skills in 31 Societies," *Social Science Research* 77 (2019): 1–15, https://doi.org/10.1016/j.ssresearch.2018.10.003.

27. M. T. Tine and A. G. Butler, "Acute Aerobic Exercise Impacts Selective Attention: An Exceptional Boost in lower-income Children," *Educational Psychology* 32, no. 7 (2012): 821–834, https://doi.org/10.1080/01443410.2012.723612.

28. H.-E. Scharfen and D. Memmert, "The Relationship Between Cognitive Functions and Sport-Specific Motor Skills in Elite Youth Soccer Players," *Frontiers in Psychology* 10 (2019): 817, https://doi.org/10.3389/fpsyg.2019.00817.

29. S. Tiberi, I. F. Marti, and M. K. Laughlin, "Cooperative Learning in Physical Education and its Effects on Student Reading Comprehension Scores," *The Physical Educator* 77, no. 2 (2020): 294–312, https://doi.org/10.18666/TPE-2020-V77-I2-9200.

30. E. A. Haapala, J. Väistö, N. Lintu, et al., "Physical Activity and Sedentary Time in Relation to Academic Achievement in Children," *Journal of Science and Medicine in Sport* 20, no. 6 (2017): 583–589, https://doi.org/10.17863/CAM.7674.

31. T. A. Uhrich and R. L. Swalm, "A Pilot Study of a Possible Effect from a Motor Task on Reading Performance," *Perceptual and Motor Skills* 104, no. 3 (2007): 1035–1041, https://doi.org/10.2466/pms.104.3.1035-104.

32. M. F. Westerveld and R. M. Armstrong, "Expository Discourse Skills of Students with Reading Comprehension Difficulties," *International Journal of Speech Language Pathology* 24, no. 6 (2022): 647–656, https://doi.org/10.1080/17549507.2022.2047784.

33. P. G. Cooker and C. A. Caffey, "Addressing the Cognitive and Affective Needs of College Athletes: Effects of Group Counseling on Self-Esteem, Reading Skills, and Coaches' Perceptions of Attitude," *Journal of Sport Psychology* 6, no. 4 (1984): 377–384, https://doi.org/10.1123/jsp.6.4.377.

34. L. A. Fuller, *Sport and the College Volunteer: Teacher Perceptions of a Supplemental Reading Program Using Social Capital Theory* (University of Connecticut, 2011), https://digitalcommons.lib.uconn.edu/gs_theses/216.

35. M. Adelantado-Renau, D. Moliner-Urdiales, I. Cavero-Redondo, M. R. Beltran-Valls, V. Martínez-Vizcaíno, and C. Álvarez-Bueno, "Association Between Screen Media Use and Academic Performance Among Children and Adolescents: A Systematic Review and meta-analysis," *JAMA Pediatrics* 173, no. 11 (2019): 1058–1067, https://doi.org/10.1001/jamapediatrics.2019.3176.

36. J. Coiro, "Predicting Reading Comprehension on the Internet," *Journal of Literacy Research* 43, no. 4 (2011): 352–392, https://doi.org/10.1177/1086296X11421979.

37. K. H. Turner, T. Hicks, and L. Zucker, "Connected Reading: A Framework for Understanding How Adolescents Encounter, Evaluate, and Engage With Texts in the Dgital Age," *Reading Research Quarterly* 55, no. 2 (2019): 291–309, https://doi.org/10.1002/rrq.271.

38. P. Panyasai, "Enhancing Reading-Comprehension Abilities and Attitudes of EFL Students Through Utilising Content-Creation Tools in Classroom Presentations," *International Journal of Learning, Teaching and Educational Research* 22, no. 7 (2023): 497–516, https://doi.org/10.26803/ijlter.22.7.26.

39. M. Courtney, M. Karakus, Z. Ersozlu, and K. Nurumov, "The Influence of ICT Use and Related Attitudes on Students' Math and Science Performance: Multilevel Analyses of the Last Decade's PISA Surveys," *Large-scale Assessments in Education* 10, no. 1 (2022): 8, https://doi.org/10.1186/s40536-022-00128-6.

40. D. Martínez-Gautier, R. Garrido-Yserte, and M. Gallo-Rivera, "Educational Performance and Icts: Availability, Use, Misuse and Context," *Journal of Business Research* 135 (2021): 173–182, https://doi.org/10.1016/j.jbusres.2021.06.027.

41. O. Navarro-Martinez and B. Peña-Acuña, "Technology Usage and Academic Performance in the Pisa 2018 Report," *Journal of New Approaches in Educational Research* 11 (2022): 130–145, https://doi.org/10.7821/naer.2022.1.735.

42. J. R. Rico-Juan, C. Cachero, and H. Macià, "Study Regarding the Influence of a Student's Personality and an LMS Usage Profile on Learning Performance Using Machine Learning Techniques," *Applied Intelligence* 54, no. 8 (2024): 1–23, https://doi.org/10.1007/s10489-024-05483-1.

43. T. Goh, B. Yang, Y. Yan, and M. Liu, "The Complex Influence of Information Communication Technology (ICT) on Students' Performance in Singapore: Evidence from 2018 PISA Data," *Computers in the Schools* 43, no. 1 (2024): 84–104, https://doi.org/10.1080/07380569.2024.2344022.

44. OECD, *Managing Screen Time: How to Protect and Equip Students Against Distraction (PISA in Focus No. 124)* (OECD Publishing, 2024), https://doi.org/10.1787/7c225af4-en.

45. G. Johnson, "Technology Use and Reading Comprehension Among Australian Indigenous Adolescents," *International Journal of Economy, Management and Social Sciences* 2, no. 8 (2013): 558–564.

46. T. Y. Tai, H. H. J. Chen, and G. Todd, "The Impact of a Virtual Reality App on Adolescent EFL Learners' Vocabulary Learning," *Computer Assisted Language Learning* 35, no. 4 (2022): 892–917, https://doi.org/10.1080/09588221.2020.1752735.

47. L. López-Agudo, M. De Guevara Rodríguez, and Ó. Marcenaro-Gutierrez, "The Influence of Grade Retention on Students' Competences in Spain," *European Journal of Education* 59, no. 4 (2024): 1–21, https://doi.org/10.1111/ejed.12736.

48. A. Constante-Amores, D. Arroyo-Resino, M. Sánchez-Munilla, and I. Asensio-Muñoz, "Contribution of Machine Learning to the Analysis of Grade Repetition in Spain: A Study Based on PISA Data," *Revista Espanola de Pedagogia* 82, no. 289 (2024): 539–562, https://doi.org/10.22550/2174-0909.4014.

49. Q. Lian, C. Yu, X. Tu, et al., "Grade Repetition and Bullying Victimization in Adolescents: A Global cross-sectional Study of the Program for International Student Assessment (PISA) Data from 2018," *PLoS Medicine* 18, no. 6 (2021): e1003846, https://doi.org/10.1371/journal.pmed.1003846.

50. H. Marsh, R. Pekrun, P. Parker, et al., "Long-Term Positive Effects of Repeating a Year in School: Six-Year Longitudinal Study of Self-Beliefs, Anxiety, Social Relations, School Grades, and Test Scores," *Journal of Educational Psychology* 109, no. 3 (2017): 425–438, https://doi.org/10.1037/edu0000144.

51. S. R. Jimerson, "Meta-Analysis of Grade Retention Research: Implications for Practice in the 21st Century," *School Psychology Review* 30, no. 3 (2001): 420–437, https://doi.org/10.1080/02796015.2001.12086124.

52. M. S. Demagistri, "Comprensión Y Producción Lectora En Adolescentes. Análisis Comparativo De Dos Franjas Etarias," in *Libro De*

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Actas Del Primer Encuentro De Grupos De Investigación Sobre Procesamiento Del Lenguaje, ed. V. J. y Y. Sevilla (Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires, 2011), 111–234.

53. P. Sainio, K. Eklund, R. Hirvonen, T. Ahonen, and N. Kiuru, "Adolescents' Academic Emotions and Academic Achievement Across the Transition to Lower Secondary School: The Role of Learning Difficulties," *Scandinavian Journal of Educational Research* 65, no. 3 (2020): 385–403, https://doi.org/10.1080/00313831.2019.1705900.

54. S. Klauda and J. Guthrie, "Comparing Relations of Motivation, Engagement, and Achievement Among Struggling and Advanced Adolescent Readers," *Reading and Writing* 28, no. 2 (2015): 239–269, https://doi.org/10.1007/S11145-014-9523-2.

55. L. Mothes, C. Kristensen, R. Oliveira, I. Argimon, R. Fonseca, and T. Irigaray, "Stressful Events and Executive Functioning in Adolescents with and Without History of Grade Repetition," *Universitas Psychologica* 16, no. 4 (2017): 1–12, https://doi.org/10.11144/JAVERIANA.UPSY16-4.SEEF.

56. A. Dhini, I. Surjandari, B. Kusumoputro, and A. Kusiak, "Extreme Learning machine–radial Basis Function (Elm-rbf) Networks for Diagnosing Faults in a Steam Turbine," *Journal of Industrial and Production Engineering* 39, no. 7 (2021): 1–9, https://doi.org/10.1080/21681015.2021.1887948.

57. J. Llorca-Schenk, J. R. Rico-Juan, and M. Sanchez-Lozano, "Designing Porthole Aluminium Extrusion Dies on the Basis of Explainable Artificial Intelligence," *Expert Systems with Applications* 222 (2023): 119808, https://doi.org/10.1016/j.eswa.2023.119808.

58. B. Liu, "New Technology Application in the Logistics Industry Based on Machine Learning and Embedded Network," *Microprocessors and Microsystems* (2020): 103596, https://doi.org/10.1016/j.micpro.2020.103484.

59. S. S. M. Ajibade, G. N. Alhassan, A. Zaidi, et al., "Evolution of Machine Learning Applications in Medical and Healthcare Analytics Research: A Bibliometric Analysis," *Intelligent Systems with Applications* 24 (2024): 200441, https://doi.org/10.1016/j.iswa.2024.200441.

60. F. Krones, U. Marikkar, G. Parsons, A. Szmul, and A. Mahdi, "Review of Multimodal Machine Learning Approaches in Healthcare," *Information Fusion* 114 (2025): 102690, https://doi.org/10.1016/j.inffus.2024.102690.

61. C. Cachero, J. R. Rico-Juan, and H. Macià, "Influence of Personality and Modality on Peer Assessment Evaluation Perceptions Using Machine Learning Techniques," *Expert Systems with Applications* 213 (2023): 1–16, https://doi.org/10.1016/j.eswa.2022.119150.

62. J. R. Rico-Juan, A.-J. Gallego, and J. Calvo-Zaragoza, "Automatic Detection of Inconsistencies Between Numerical Scores and Textual Feedback in peer-assessment Processes with Machine Learning," *Computers & Education* 140 (2019): 1–13, https://doi.org/10.1016/j.compedu.2019.103609.

63. M. Alvarez-Garcia, M. Arenas-Parra, and R. Ibar-Alonso, "Uncovering Student Profiles. an Explainable Cluster Analysis Approach to PISA 2022," *Computers & Education* 223 (2024): 105166, https://doi.org/10.1016/j.compedu.2024.105166.

64. P. W. Koh and P. Liang, "Understanding black-box Predictions via Influence Functions," *Proceedings of the 34th International Conference on Machine Learning, ICML'17* (2017): 1885–1894.

65. D. Liang, D. A. Frederick, E. E. Lledo, et al., "Examining the Utility of Nonlinear Machine Learning Approaches Versus Linear Regression for Predicting Body Image Outcomes: The U.S. Body Project I," *Body Image* 41 (2022): 32–45, https://doi.org/10.1016/j.bodyim.2022.01.013.

66. A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, et al., "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI," *Information Fusion* 58 (2020): 82–115, https://doi.org/10.1016/j.inffus.2019.12.012.

67. S. Lundberg, "Shap (Shapley Additive Explanations)," (2019), https://github.com/slundberg/shap.

68. B. Tan and M. Cutumisu, "Employing Tree-based Algorithms to Predict Students' Self-Efficacy in PISA 2018," in *Proceedings of the 15th International Conference on Educational Data Mining* (2022), 1–6, https://educationaldatamining.org/EDM2022/proceedings/2022.EDM-posters.73/.

69. X. Miao, A. Nadaf, and Z. Zhou, "Machine Learning Evidence from PISA 2018 Data to Integrate Global Competence Intervention in UAE K–12 Public Schools," *International Review of Education* 69, no. 5 (2023): 675–690, https://doi.org/10.1007/s11159-023-10026-w.

70. A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics* 12, no. 1 (1970): 55–67, https://doi.org/10.1080/00401706.1970.10488634.

71. R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B* 58, no. 1 (1996): 267–288, https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

72. D. J. MacKay, "Bayesian Nonlinear Modeling for the Prediction Competition," *ASHRAE transactions* 100, no. 2 (1994): 1053–1062.

73. L. Breiman, *Classification and Regression Trees* (Routledge, 2017), https://doi.org/10.1201/9781315139470.

74. L. Breiman, "Random Forests," *Machine Learning* 45, no. 1 (2001): 5–32, https://doi.org/10.1023/A:1010933404324.

75. Y. Freund and R. E. Schapire, "A decision-theoretic Generalization of on-line Learning and an Application to Boosting," *Journal of Computer and System Sciences* 55, no. 1 (1997): 119–139, https://doi.org/10.1006/jcss.1997.1504.

76. T. Chen and C. Guestrin, "Xgboost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2016* (ACM, 2016), 785–794, https://doi.org/10.1145/2939672.2939785.

77. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: Unbiased Boosting with Categorical Features," *Advances in Neural Information Processing Systems* 31 (2018).

78. G. Ke, Q. Meng, T. Finley, et al., "Lightgbm: A Highly Efficient Gradient Boosting Decision Tree," *Advances in Neural Information Processing Systems* 30 (2017): 3146–3154.

79. F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-Learn: Machine Learning in Python," *Journal of Machine Learning Research* 12 (2011): 2825–2830.

80. G. E. Hinton, "Connectionist Learning Procedures," *Artificial Intelligence* 40, no. 1-3 (1989): 185–234, https://doi.org/10.1016/0004-3702(89)90049-0.

81. A. Kassambara, *Machine Learning Essentials: Practical Guide in R. Sthda* (2018).

82. S. Chatterjee and A. S. Hadi, *Regression Analysis by Example* (John Wiley & Sons, 2015).

83. S. Van Buuren and K. Groothuis-Oudshoorn, "Mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software* 45, no. 3 (2011): 1–67, https://doi.org/10.18637/jss.v045.i03.

84. T. A. Warm, "Weighted Likelihood Estimation of Ability in Item Response Theory," *Psychometrika* 54, no. 3 (1989): 427–450, https://doi.org/10.1007/BF02294627.

85. A. E. Roth, *The Shapley Value: Essays in Honor of Lloyd S* (Shapley. Cambridge University Press, 1988), https://doi.org/10.1017/CBO9780511528446.

86. S. M. Lundberg and S.-I. Lee, "Consistent Feature Attribution for Tree Ensembles," *arXiv* (2017): 1–7, https://doi.org/10.48550/arXiv.1706.06060.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Cbua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

87. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, ed. M. I. Jordan, Y. LeCun, and S. A. Solla (The MIT Press, 2017), 4765–4774.

88. F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometric Bulletin* 1, no. 6 (1945): 80–83, https://doi.org/10.2307/3001968.

89. K. Stapor, P. Ksieniewicz, S. García, and M. Woźniak, "How to Design the Fair Experimental Classifier Evaluation," *Applied Soft Computing* 104 (2021): 107219, https://doi.org/10.1016/j.asoc.2021.107219.

90. S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics* (1979): 65–70, https://www.jstor.org/stable/4615733.

91. L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A Novel Bandit-based Approach to Hyperparameter Optimization," *Journal of Machine Learning Research* 18, no. 1 (2017): 6765–6816.

92. K. Pearson, "X. on the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it can be Reasonably Supposed to Have Arisen from Random Sampling," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* 50, no. 302 (1900): 157–175, https://doi.org/10.1080/14786440009463897.

93. V. Chernozhukov, D. Chetverikov, M. Demirer, et al., *Double/Debiased Machine Learning for Treatment and Structural Parameters* (2018), https://doi.org/10.1111/ectj.12097.

94. H. B. Mann and D. R. Whitney, "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics* 18, no. 1 (1947): 50–60, https://doi.org/10.1214/aoms/1177730491.

95. B. Klimova and K. Zamborova, "Use of Mobile Applications in Developing Reading Comprehension in Second Language acquisition—A Review Study," *Education Sciences* 10, no. 12 (2020): 391, https://doi.org/10.3390/educsci10120391.

96. M. Vásquez-Carbonell, "A Systematic Literature Review of Educational Apps: What Are They up To?" *Journal of Mobile Multimedia* 18, no. 2 (2021): 251–274, https://doi.org/10.13052/jmm1550-4646.1825.

97. M. Beek, M. Opdenakker, A. Spijkerboer, L. Brummer, H. Ozinga, and J. Strijbos, "Scaffolding Expository History Text Reading: Effects on Adolescents' Comprehension, self-regulation, and Motivation," *Learning and Individual Differences* (2019): https://doi.org/10.1016/J.LINDIF.2019.06.003.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Supporting Information.** SHAP values for all predictors of reading comprehension.

hbet, 2026, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1155/hbe2/8388818 by Chua - Universidad De Huelva, Wiley Online Library on [28/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License