rep

revista española
de pedagogía

n. 289

Revista Española de Pedagogía
year 82, n. 289, September-December 2024
Revista Española de Pedagogía
año 82, n.º 289, septiembre-diciembre 2024

Table of contents
Sumario

Articles
Artículos

Book reviews

Table of contents of the year 2024

Instructions for authors

Articles

Ahmed Tlili

Can artificial intelligence (AI) help in computer science education?
A meta-analysis approach

Leidy-Vanessa Díaz-Beltrán and Pablo Páramo

Teaching in pro-environmental rules:
Implications for environmental education and climate change

Francisco López-Rupérez, Álvaro Moraleda-Ruano and Isabel García-García

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

**Alexander Constante-Amores, Delia Arroyo-Resino, María Sánchez-Munilla
and Inmaculada Asensio-Muñoz**

Contribution of machine learning to the analysis of grade repetition in Spain:
A study based on PISA data

Dragana Pavlovic, Joan Soler-Adillon and Zorica Stanisavljevic-Petrovic

A full-time professor dedicated to you: ChatGPT from university students' perspective

Alicia Palacios, Rosa Gómez, Álvaro Barreras and Daniel Moreno-Mediavilla

Teachers' attitudes towards the use of virtual simulations:
Design and validation of a questionnaire

**María-José Mera-Lemp, Javier Torres-Vallejos
and Florencia Guglielmetti-Serrano**

Chilean teachers' cultural threat and attitudes towards multiculturality at school:
The role of outgroup anxiety and intercultural sensitivity

Can artificial intelligence (AI) help in computer science education? A meta-analysis approach

¿Puede ayudar la inteligencia artificial (IA) en la educación en ciencias de la computación? Un enfoque metaanalítico

Ahmed TLILI. .Associate Professor. Smart Learning Institute of Beijing Normal University, China (*ahmedtlili@ieee.org*).

Abstract:

Several studies have investigated the effect of Artificial Intelligence (AI) on students' learning achievement in education. However, limited research targeted Computer Science (CS) education, which is considered crucial regardless of the future profession. Consequently, scant information exists on how AI might impact students' learning achievement in CS education. To address this research gap, this study conducts a systematic review and a meta-analysis to investigate how AI integration affects learning achievement in CS education and the potential moderating variables of this effect. Specifically, 28 studies (n = 2765 participants in total) were included and meta-analyzed, and the obtained effect size was very large (g = 1.36, p <.001). Particularly, intelligent tutoring systems (ITSs) were found to have the highest effect (g = 1.45; huge) as an AI technology. Additionally, the AI intervention duration and the geographical distribution of students are found to moderate the AI effect in CS education. The findings of this study can serve as a reference for various stakeholders (e. g., educators, computer scientists, instructional designers) on how to integrate AI and improve learning experiences and outcomes in CS education.

Keywords: computer science, computing, artificial intelligence, education, learning, collaborative intelligence, meta-analysis, learning achievement.

Resumen:

Varios estudios han investigado el efecto de la inteligencia artificial (IA) en los logros de aprendizaje de los alumnos

en educación. Sin embargo, hay un volumen limitado de investigación enfocada en la educación en ciencias informáticas o de la computación (CC), un área que se considera crucial con independencia de la profesión futura. En consecuencia, existe poca información sobre cómo la IA podría influir en los logros de aprendizaje de los alumnos en la educación en CC. A fin de llenar este vacío en la literatura, este estudio realiza una revisión sistemática y un metaanálisis para investigar cómo la integración de la IA afecta a los logros de aprendizaje en la educación en CC y las posibles variables moderadoras de este efecto. En concreto, se incluyeron y metaanalizaron 28 estudios (n = 2765 participantes en total) y el tamaño del efecto obtenido fue muy grande (g = 1.36, p <0.001). En particular, se ha hallado que los sistemas de tutoría inteligente (STI) presentan el mayor efecto (g = 1.45, un efecto enorme) como tecnología de IA. Adicionalmente, se ha encontrado que la duración de la intervención de la IA y la distribución geográfica de los alumnos moderan el efecto de la IA en la educación en CC. Los hallazgos de este estudio pueden servir como referencia para las diversas partes interesadas (por ejemplo, educadores, científicos computacionales y diseñadores formativos) sobre cómo integrar la IA y mejorar las experiencias y los resultados de aprendizaje en la educación en CC.

1. Introduction

1.1. Artificial intelligence in computer science education

The increased need to educate students about computer science (CS) has been highlighted by former U.S President Obama, who initiated "CSforAll" to provide all K–12 children CS education in the United States (Smith, 2016). After that, several initiatives have been launched to teach CS or a particular technology, such as Artificial Intelligence (AI), to all students. For instance, the Computer Science Teachers Association (CSTA) and the Association for the Advancement of Artificial Intelligence (AAAI) announced in 2018 the establishment of national guidelines about teaching K-12 students AI. These guidelines revolve around "five big ideas" in AI, namely Perception, Representation & Reasoning, Learning, Natural Interaction, and Societal Impact (AAAI, 2018; Touretzky et al., 2023). Teaching students CS and the use of technology is becoming a crucial competence for everyday life and any future profession.

However, learning CS subjects might be difficult and require more cognitive tasks than other subjects (Tlili et al., 2015). Silva et al. (2019) also pointed

out the challenges of teaching CS when analyzing the different disciplines in the area. Particularly, the difficulties that students face when learning about algorithms and data structure are of major concern for educators worldwide (Silva et al., 2019). In the same vein, Waraich (2004) further highlighted that students find some computer science subjects like binary arithmetic and logic gates to be "dry" and "not very interesting".

AI, on the other hand, has also been reported to support education by providing personalized learning content, recommendation, and predictions (Onesi-Ozigagun et al., 2024). This has triggered researchers to integrate AI as a technology to facilitate CS education. For instance, Georgia Institute of Technology integrated AI-based instructions to teach its online master students about AI, which is part of a computer science program (Goel & Joyner, 2017). Zhang et al. (2021) developed an AI-based interactive e-book to teach K-12 students about AI. Their findings revealed that students have a high acceptance level of this e-book for learning. Gerdes et al. (2017) developed an adaptable programming tutor for learning the Haskell programming language.

1.2. Research gap and study objectives

Despite the increase adoption of AI in education generally and in CS particularly, the effects of AI on students' learning achievement are still debatable (Kim & Lee, 2023). Several meta-analy-

sis studies were conducted in the literature to investigate the effects of AI on learning achievement (Lin et al., 2022; Zheng et al., 2023). However, these meta-analyses targeted education in general. Particularly, scant information exists on the effect of AI on students' learning achievement in CS education. In this context, several studies conducted a systematic review of the use of Intelligent Tutoring Systems (ITS) in CS education (Crow et al., 2018; Francisco & de Oliveira Silva, 2022). However, these studies only conducted a qualitative analysis and did not discuss learning achievement. Therefore, the impact of AI on learning achievement remains unknown.

One study by Nesbit et al. (2014) conducted a meta-analysis of using ITSs in CS education, where 22 articles published between 1998 and 2013 were analyzed. The overall effect size was moderate (g = .46). While this study (Nesbit et al., 2014) is considered one of the early attempts to discuss the present research topic, its results cannot be generalized because: of two reasons. First, it focused only on one type of AI technology (i. e., ITS) rather than covering all AI technologies. Therefore, the obtained results do not reflect the AI effect in general. Second, the articles covered were before 2013 (with one study in 2013), hence the findings can be considered outdated due to the rapid evolution of technology generally and AI particularly, where new waves are seen like generative AI (Tlili et al., 2023), among others.

rep

To the best of our knowledge, no research conducted an analysis to investigate the effects of AI on learning achievement in CS education. Therefore, the question of whether AI as computer science technology can help in CS education remains open. While there is no agreed-upon definition of computer science, this study considers CS as the study of computers and of the phenomena connected with computing, notably algorithms, programs, and programming (Shaw, 1985). To address the highlighted research gap, this study answers the following research questions:

RQ1. What is the effect of AI on students' learning achievement in computer science education?

RQ2. Which variables moderate the effect of AI on students' learning achievement in computer science education?

2. Methodology

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines were followed (Page et al., 2021) to conduct this meta-analysis.

2.1. Data selection

A search was conducted in the following electronic databases: IEEE Xplore, Science Direct, Scopus, Taylor & Francis, and Web of Science. These databases were chosen because they are popular in the field of educational technology. The search timeframe was set starting from 2011 because: (1) this year was considered the year where AI applications became more mature (Wang et al., 2023); and (2) since then, the research on the use of AI technologies in CS education has started increasing significantly (Nesbit et al., 2015). The search keywords were adapted from several AI in education reviews in the literature (e.g., Nesbit et al., 2015; Zawacki-Richter et al., 2019; Zheng et al., 2023) and are as follows: ((Artificial intelligence substring) AND (Computer science substring)), where:

- Artificial intelligence substring: artificial intelligence OR AI OR machine intelligence OR machine learning OR natural language processing OR deep learning OR robotic.

- Computer science substring: computer science OR technology OR computer literacy OR information and communication technology OR ICT.

All publication types (conference proceedings, journal articles, etc.) were considered because proceeding papers are the most common format for researchers in computer science and software engineering, unlike in the education field, where journal publication is more common (Nesbit et al., 2015).

The final search was conducted on January 1, 2024, and the whole process yielded 548 potential studies. After

removing duplicates, 205 potential studies were identified and put through the inclusion/exclusion criteria. A study was included if it (1) was in English, (2) was empirical research, (3) used AI for CS education, (4) was not qualitative or review research, (5) provided sufficient information (e. g., mean, median, and standard deviation) to calculate the effect size, and (6) included a control condition. Finally, 28 studies (2765 participants in total) were considered for this meta-analysis.

2.2. Meta-analysis

Comprehensive Meta-Analysis V.4 (Borenstein, 2022) software was used to conduct the current meta-analysis. In addition, Hedges' g was used to calculate the effect sizes (Hedges, 1981). The motivation behind using Hedges' g instead of Cohen's d effect size was that the differential sample size between studies may bias the estimated effect size. This bias affects studies with a sample size smaller than 20, in which case Hedges's g presents more reliable estimates than Cohen's d (Hedges & Olkin, 1985). Table 1 presents the 28 included studies in this present meta-analysis.

Three methods were used to assess publication bias. Firstly, the trim-and-fill method, with the intention of identifying publication bias by means of a funnel plot wherein the studies are represented by dots. If the dots are distributed on both sides of a vertical line, representing the average effect size, it is assumed that there is no publication bias (Borenstein et al., 2010). Secondly, Rosenthal's (1979) fail-safe number aims to determine the number of studies with nonsignificant results of unpublished data needed to nullify the mean effect size. A fail-safe number larger than 5k+10 (where k is the original number of studies included in the meta-analysis) is robust. This means that the effect size of unpublished studies is not likely to affect the average effect size of the meta-analysis. However, this method assumes that the mean effect size in the missing studies is zero (Borenstein et al., 2021). The third method was Egger's regression test, where a significant intercept suggests publication bias (Lin et al., 2018).

2.3. Coding scheme

To minimize the potential for bias, an online electronic data extraction form was designed (Kitchenham & Charters, 2007). This form contains the following information, which could help in answering the afromentioned research questions: (1) AI technology: it refers to the type of AI technology used during the learning process; (2) field of education: it refers to the educational domain where AI was used; (3) level of education: it refers to the academic level at which AI was used; (4) learning mode: it refers to where and how the learning process occurred; (5) intervention duration: it refers to the time length over which AI was used in education; and (6) geographical distribution: it refers to the geographical distribution of students who were involved in the learning process.

Table 1. Descriptive and inferential statistics of the 28 included studies.

#	Author(s) & year	AI type	Level of education	Learning mode	Intervention duration	Geographical distribution	Hedges's g	Standard error	Sample		
									Number of experimental	Number of control	Total
1	Hayashi (2019)	ITS	Primary	Online	One month	Asia	1.335	0.362	18	18	36
2	Cuong et al. (2018)	ITS	Higher education	Online	One semester	Asia	1.297	0.255	37	36	73
3	Yin et al. (2021)	Chatbot	Higher education	Online	One semester	Asia	0.000	0.200	51	48	99
4	Perikoset (2017)	ITS	Higher education	Online	One month	Europe	3.381	0.207	113	113	226
5	Tegos and Demetriadis (2017)	Chatbot	Higher education	Online	One semester	Europe	0.859	0.244	38	34	72
6	Tegos et al. (2015)	Chatbot	Higher education	Online	One week	Europe	0.903	0.315	21	22	43
7	Tegos et al. (2016)	Chatbot	Higher education	Online	One week	Europe	0.594	0.252	32	32	64
8	Fidan and Gencel (2022)	Chatbot	Higher education	Online	One month	Europe	9.320	0.711	54	40	94
9	Abbasi et al. (2019)	Chatbot	Higher education	Online	One semester	Asia	1.214	0.206	55	55	110
10	Abbasi and Kazi (2014)	Chatbot	Higher education	Online	One semester	Asia	1.852	0.280	36	36	72
11	Grivokostopoulou et al. (2017)	ITS	Higher education	Online	One month	Europe	1.929	0.140	150	150	300
12	Essel et al. (2022)	Chatbot	Higher education	Online	One semester	Africa	5.592	0.536	34	34	68
13	Lin and Chen (2020)	Personalized	Higher education	Blended	One month	Asia	1.337	0.223	48	49	97

#	Author(s) & year	AI type	Level of education	Learning mode	Intervention duration	Geographical distribution	Hedges's g	Standard error	Sample		
									Number of experimental	Number of control	Total
14	Stamper et al. (2013)	ITS	Higher education	Online	One semester	North America	0.086	0.140	105	98	203
15	Ma et al. (2020)	ITS	Secondary	Online	One month	Asia	4.777	0.455	38	36	74
16	Su (2020)	ITS	Higher education	Online	One month	Asia	0.598	0.226	40	40	80
17	Chang et al. (2016)	ITS	Higher education	Blended	One semester	Asia	0.552	0.188	58	58	116
18	Lai et al. (2021)	Personalized	Higher education	Online	One semester	Asia	1.726	0.254	43	41	84
19	Eryilmaz and Adabashi (2020)	ITS	Higher education	Online	One semester	Europe	0.342	0.183	60	60	120
20	Arsovic and Stefanovic (2020)	Personalized	Higher education	Online	One semester	Europe	0.628	0.161	81	79	160
21	Winkler et al. (2020)	Chatbot	Higher education	Online	One semester	Europe	0.682	0.240	37	35	72
22	Song and Kim (2021)	Chatbot	Higher education	Online	One semester	North America	0.553	0.269	27	29	56
23	Kumar (2021)	Chatbot	Higher education	Online	One semester	Asia	0.243	0.256	30	30	60
24	Farah et al. (2022)	Chatbot	Higher education	Online	One month	Europe	0.205	0.432	11	9	20
25	Al-Abdullatif et al. (2023)	Chatbot	Higher education	Blended	One semester	Asia	0.338	0.257	30	30	60
26	Ortega-Ochoa et al. (2024)	Chatbot	Higher education	Online	One semester	Europe	0.243	0.143	101	95	196
27	Sun et al. (2024)	Chatbot	Higher education	Online	One semester	Asia	0.306	0.220	43	39	82
28	Ahn and Oh (2024)	ITS	Higher education	Blended	One semester	Asia	0.459	0.314	20	20	40

3. Results

3.1. Publication bias assessemnt

Figure 1 shows that the dots in this study are almost distributed symmetrically around the vertical line. Additionally, although a few dots are outside the triangle of the funnel plot, they are in the upper part of Figure 1 and not at the bottom. Borenstein et al. (2010) stated that a symmetric funnel plot implies that there is no publication bias. Therefore, it is concluded that the meta-analysis reliability of this study is not affected by publication bias.

FIGURE 1. Funnel plot of standard error by Hedges's *g*.

Funnel plot of standard error by Hedges' *g*
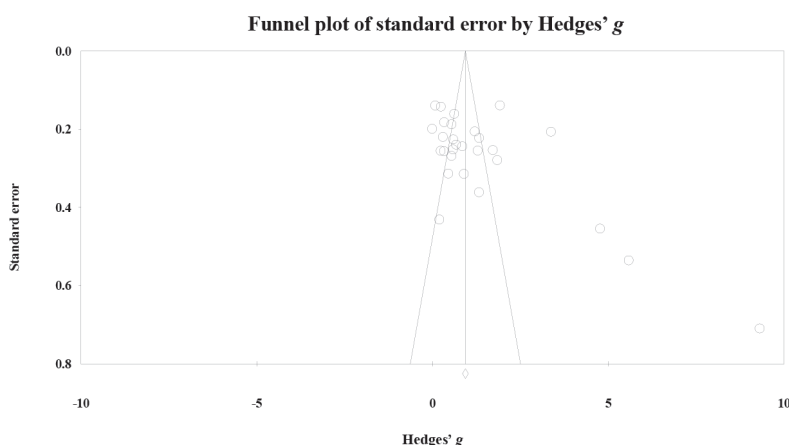
3.2. Effect of AI on students' learning achievement in CS education

Figure 2 presents the forest plot of the variation of effect size across the 28 included papers. The black square represents each paper's weighted effect size, where a larger square size implies a larger effect size. The arrow underneath each square (effect size) represents the confidence interval of the associated effect size. The overall mean effect size (g = 1.364) is presented in the last row of the forest plot. Overall, most papers had a positive effect size with different confidence intervals.

Table 2 shows that the meta-analysis yielded an overall effect size of g = 1.36, p <.001, indicating that AI had a very large effect on students' learning achievement in CS education. Particularly, it is seen that all AI technologies had a significant positive effect, with some variation regarding the effect size; chatbots (g = 1.35, 95% CI = 0.77 to 1.94) and personalized learning systems (g = 1.21, 95% CI = 0.54 to 1.87) had a very large effect on students' learning achievement, while ITS (g = 1.45, 95% CI = 0.68 to 2.21) had a huge effect.

The I^2 statistic showed that 95.77% of variance resulted from between-study factors, implying that other variables might moderate the AI effect size on students' learning achievement. It is crucial, therefore, to conduct further analysis and investigate the potential moderating variables.

FIGURE 2. Forest plot of the Hedge's g estimates and the confidence intervals

Study name	Subgroup within study	Hedges' g	Standard error	Hedges' g and 95% CI
Al-Abdullatif et al. (2023)	Chatbot	0.338	0.257	
Hayashi, Y. (2019)	ITS	1.335	0.362	
Cuong et al(2018)	ITS	1.297	0.255	
Yin et al. (2021)	Chatbot	0.000	0.200	
Perikoset (2017)	ITS	3.381	0.207	
Tegos, & Demetriadis (2017)	Chatbot	0.859	0.244	
Tegos et al.(2015)	Chatbot	0.903	0.315	
Tegos et al.(2016)	Chatbot	0.594	0.252	
Fidan, & Gencel (2022)	Chatbot	9.320	0.711	
Abbasi et al(2019)	Chatbot	1.214	0.206	
Abbasi & Kazi (2014)	Chatbot	1.852	0.280	
Grivokostopoulou et al(2017)	ITS	1.929	0.140	
Essel et al(2022)	Chatbot	5.592	0.536	
Lin, & Chen (2020)	Personalized	1.337	0.223	
Stamper et al(2013)	ITS	0.086	0.140	
Ma et al(2020)	ITS	4.777	0.455	
Su (2020)	ITS	0.598	0.226	
Chang et al., (2016)	ITS	0.552	0.188	
Lai, Y. H et al(2021)	Personalized	1.726	0.254	
Ery?lmaz, & Adabashi, (2020)	ITS	0.342	0.183	
Arsovic & Stefanovic (2020)	Personalized	0.628	0.161	
Winkler et al(2020)	Chatbot	0.682	0.240	
Song & Kim (2021)	Chatbot	0.553	0.269	
Kumar (2021)	Chatbot	0.243	0.256	
Farah et al(2022)	Chatbot	0.205	0.432	
Ortega-Ochoa et al., (2024)	Chatbot	0.249	0.143	
Sun et al., (2024)	Chatbot	0.306	0.220	
Ahn & Oh (2024)	ITS	0.459	0.314	
		1.364	0.210	

TABLE 2. Effect of AI on learning achievement in computer science education.

Analysis	n	g	95 % CI	Z	p	I^2	τ^2	Effect size interpretation
Overall	28	1.36	[0.95, 1.78]	6.50	0.001***	95.77	1.15	Very large
Chatbot	15	1.35	[0.77, 1.94]	4.54	0.001***	95.12	1.23	Very large
ITS	10	1.45	[0.68, 2.21]	3.71	0.001***	97.09	1.45	Huge
Person-alized systems	3	1.21	[0.54, 1.87]	3.56	0.001***	87.18	0.30	Very large

Note: n = number of studies; g = Hedges' g effect size; CI = confidence interval; Z = Z value for Hedges' g; p = p values of Hedges' g; I^2 and τ^2 are measures of effect size variability; ***p <.001.

3.3. Moderating variables of AI effect in CS education

To investigate the moderating variables of the AI effect, this section starts by first investigating how the AI effect varies across various variables, including educational level (see Table 3), learning mode (see Table 4), AI intervention duration (see

Table 5), and geographical distribution of students (see Table 6). It then runs a meta-regression model taking into consideration all of the mentioned variables (see Table 7) to identify the moderating variables of the AI effect in CS education.

Table 3 presents the effect of AI on students' learning achievement across var-ious educational levels. It is seen that AI was used mostly in higher education ($n = 26$) to teach about computer science sub-jects. The effect was very large ($g = 1.24$). On the other hand, one study for each educational level, namely primary and secondary education, found that the ef-fects were very large ($g = 1.35$) and huge ($g = 4.77$), respectively.

TABLE 3. Effect of AI on learning achievement across various educational levels.

Level of education	n	g	95 % CI	Z	p	I^2	τ^2	Effect size interpretation
PE	1	1.35	[0.63, 2.04]	3.69	0.001***	0	0	Very large
SE	1	4.77	[3.89, 5.67]	10.50	0.001***	0	0	Huge
HE	26	1.24	[0.83, 1.65]	5.98	0.001***	95.58	1.04	Very large

Note: n = number of studies; g = Hedges' g effect size; CI = confidence interval; Z = Z value for Hed-ges' g; p = p values of Hedges' g; I^2 and τ^2 are measures of effect size variability; ***p <.001; *p <0.5.

TABLE 4. Effect of AI on learning achievement in different learning modes.

Modo de aprendizaje	n	g	95 % CI	Z	p	I^2	τ^2	Effect size interpretation
Blended	4	0.69	[0.23, 1.14]	2.94	0.003**	75.53	0.16	Medium
Online	24	1.49	[1.02, 1.97]	6.91	0.001***	96.31	1.31	Huge

Note: n = number of studies; g = Hedges' g effect size; CI = confidence interval; Z = Z value for Hedges' g; p = p values of Hedges' g; I^2 and τ^2 are measures of effect size variability; ***p <.001.

The effect of AI in different learning modes on students' learning achievement was analyzed (see Table 4). It is evident that AI has a significant positive effect on learning achievement in all learning modes. Specifically, the application of AI in blended learning had a medium effect ($g = 0.69$), while the effect was very large in online learning ($g = 1.49$).

The effect of AI with different inter-vention durations on students' learning achievement was analyzed (see Table 5). It is seen that AI, regardless of the intervention duration, had a significant positive effect on learning achievement in CS education. Specifically, the effect of AI was huge ($g = 2.74$) when it was used between one week and one month.

However, this effect decreases (to be large) if AI is used for less or more than that period.

The effect of AI on students' learning achievement, based on their geographical distribution, was analyzed (see Table 6). It is seen that AI had a significant positive effect across all continents. Specifically, AI has a huge effect on the learning achievement of students in Africa (g = 5.59), while the effect was very large in Asia (g = 1.10) and Europe (g = 1.59). Finally, the effect was small in North America (g = 0.26).

To further investigate for possible covariance between the analyzed variables above, a meta-regression that includes all moderators was conducted. The obtained results revealed that the effect of AI on learning achievement in CS education significantly varies according to the AI intervention duration (Q = 6.72, df = 2, p = 0.03) and the geographical distribution of students (Q = 14.47, df = 3, p = 0.001).

TABLE 5. Effect of AI on learning achievement with different intervention durations.

Duration	n	g	95 % CI	Z	p	I^2	τ^2	Effect size interpretation
Duration < 1 week	2	0.72	[0.33, 1.10]	3.63	0.001***	0	0	Large
1 week ≤ duration < 1 month	8	2.74	[1.63, 3.84]	4.86	0.001***	97.27	2.40	Huge
1 month ≤ duration < 1 semester	18	0.85	[0.52, 1.17]	5.10	0.001***	90.53	0.44	Large

Note: n = number of studies; g = Hedges' g effect size; CI = confidence interval; Z = Z value for Hedges' g; p = p values of Hedges' g; I^2 and τ^2 are measures of effect size variability; ***p <.001; *p <.05.

TABLE 6. Effect of AI on learning achievement based the geographical distribution

Continent	n	g	95 % CI	Z	p	I^2	τ^2	Effect size interpretation
Africa	1	5.59	[4.54, 6.64]	10.43	0.001***	0	0	Huge
Asia	14	1.10	[0.65, 1.54]	4.84	0.001***	91.38	0.64	Very large
Europe	11	1.59	[0.83, 2.36]	4.08	0.001***	97.26	1.58	Very large
North America	2	0.26	[-0.18, 0.71]	1.16	0.254	57.77	0.06	Small

Note: n = number of studies; g = Hedges' g effect size; CI = confidence interval; Z = Z value for Hedges' g; p = p values of Hedges' g; I^2 and τ^2 are measures of effect size variability; ***p <.001; *p <.05.

TABLE 7. Meta-regression results for the learning achievement of response from level of education, duration, geographical distribution, and learning mode.

Model		Coefficient	Standard error	95% lower	95% upper	Z-value	2 sided p value	
Intercept		0.94	0.51	-0.07	1.94	1.83	0.07	
Level of Education	1 = PE	-1.12	1.41	-3.89	1.65	-0.79	0.43	$Q^* = 4.92$, $df = 2$, $p = 0.09$
	2 = SE	2.93	1.56	-0.14	5.99	1.87	0.06	
Duration	1 = Duration > 1 week	-1.45	1.00	-2.40	1.51	-0.45	0.66	$Q^* = 6.72$, $df = 2$, $p = 0.03^*$
	2 = 1 week ≤ duration < 1 month	1.52	0.66	0.22	2.81	2.30	0.02*	
Geographical distribution	1 = Africa	5.62	1.42	2.48	8.04	3.71	0.001***	$Q^* = 14.47$, $df = 3$, $p = 0.001^{***}$
	2 = Europe	0.25	0.62	-0.97	1.48	0.41	0.68	
	3 = North America	-0.62	0.99	-2.56	1.32	-0.63	0.53	
Learning mode	1 = Blended	-0.35	0.75	-1.82	1.12	-0.47	0.64	NA

Note ***p <.001 , *p <.05

4. Discussions

4.1. AI effect on students' learning achievement in CS education

This present study revealed that AI has a very large effect on learning achievement when it is used in computer science education compared to education generally, where several studies in the literature found that the effect was moderate (García-Martínez et al., 2023; Hwang, 2022; Wu & Yu, 2023). This could be because AI provides innovative tools and solutions that cater to the different advanced needs and competencies in computer science (Barnes et al., 2017). AI is also implemented based on a deep understanding of computer science pedagogy (Barnes et al., 2017), hence better contributing to learning achievement in CS education compared to other educational subjects/fields. Additionally, making full use of technology in education requires some competencies and a good background in understanding and using it. This requirement might be fulfilled by CS teachers who are considered to have good knowledge and expertise of AI (or technology generally) compared to other techers from other educational fields (e.g., languages, science, etc.). Particularly, ITS, as an AI technology, had the highest achievement (huge) compared to other AI technologies, namely chatbots and personalized learning systems. This could be because ITSs provide instructions and interact with students at a finer level of granularity, allowing learning to be appropriate to students' different needs and individual differences (Lin et al., 2023).

4.2. Moderating variables of the AI effect on students' learning achievement in CS education

This study revealed that educational level does not moderate the AI effect in CS education. While this present study did not focus during the search process on a specific educational level or on computer science as a curriculum (rather as educational subjects that are within CS), the obtained results revealed that AI is mostly used in higher education when teaching CS. This might be because computer science subjects are mostly taught in higher education, despite several attempts to teach some subjects like programming or AI at earlier educational levels (e. g., primary education or K–12). This could also be because students in higher education have the required competencies to use and work with AI compared to students in earlier educational levels. Particularly, the effect of AI on higher education students was very large in CS education. This might be because the integration of AI in higher education could easily adapt teaching to the various needs and profiles of learners (Verdú et al., 2017), provide individual and personalized feedback and improve assessments (Dever et al., 2020), and predict learning achievement (Çağataylı & Çelebi, 2022).

The results indicated that AI had a higher effect (a huge effect) when used to teach CS in online environments

than in blended environments (a medium effect). This could be explained by the fact that implementing collaborative intelligence (i.e., combining human intelligence and artificial intelligence) will be easier in online learning environments, as human-based and machine-based instructions can be designed, automated, and combined effectively in online learning models through the support of AI techniques and algorithms, as well as the collection and analysis of interaction (big) data of both teachers and students (i.e., learning analytics). Another reason might be because CS education requires the transfer of computational, mathematical, and algorithmic thinking (Pirker et al., 2014), which are better transferred through online instructions instead of using traditional instructions in classrooms. Although the learning mode did not moderate the effect of AI in CS education, the obtained results hold significance for educational developers and instructional designers, providing insights into potential integration of AI in future CS education and programs.

The results revealed that the AI had a huge effect on learning achievement when the intervention duration was between one week and one month. However, this effect decreased when AI was used for less or longer than this duration (between one week and one month). This might be because this duration is considered sufficient to promote students' familiarization and stimulate motivation for proper engagement and performance (Merilampi et al., 2014), while a shorter time may be insufficient to produce the desired positive outcomes (Sung et al., 2016). It is further seen that the longer the AI intervention is, the lower the effect on learning achievement. This can be explained, as highlighted by several methodologists, by the fact that longer interventions might lead to poorer implementation fidelity, hence causing a shift from the initial experiment goals and achieving low effects (Mihalic, 2004; Wang et al., 2023). The decrease of the AI effect over longer periods might also be due to the novelty effect (Pisapia et al., 1993), where learners lose interest in a given technology as time passes. The intervention duration was found to significantly moderate the AI effect in CS education, revealing the need for systematic thinking about how long AI will be used and how it can achieve the desired educational objective(s) within a specific duration.

The results indicated that the AI effect on learning achievement in computer science education varies according to students' geographical distribution and also moderates the AI effect in CS education. This is because CS education has geographic particularities that should be considered (Francisco & De Oliveira, 2022). Particularly, the results revealed that AI has a huge effect on students learning achievement in Africa. However, this result cannot be generalized, as only one study was obtained in that sub-category (see

Table 6). On the other hand, the effect of AI on learning achievement was very large in Europe and Asia. This high effect could be attributed to the long and continuous efforts of the European Commission in developing AI literacy (Tangi, 2022) and promoting the safe and responsible implementation of AI (Jarota, 2023). This has led to clear policies, rich experiences, and organized educational systems that incorporate AI, resulting in better learning outcomes. Similarly, several Asian countries are now incorporating AI into their teaching and learning programs. This might have resulted in an improvement in learning outcomes. In this context, Nye (2015) highlighted that the community of AI in education is rapidly increasing and recognizing the importance of designing technologies globally.

5. Conclusions, implications and future directions

This study conducted a meta-analysis of 28 papers to investigate the AI effect on students' learning achievement in CS education. The obtained results revealed that AI has a very large effect on learning achievement. Particularly, the AI intervention duration and the geographical distribution of students are found to moderate this AI effect in CS education.

5.1. Implications

The findings of this study can contribute to the literature from various perspectives. From a theoretical perspective, this study contributes to the ongoing debate about the effect of AI in education generally and in CS education specifically. It highlights what and how should be considered when implementing AI in CS education, hence contributing to developing frameworks and theories in this regard. Additionally, this study revealed a disparity in the AI effect within CS education across the geographical distribution of students. Therefore, more round-table discussions, policy making, and cross-country collaborations should be established to increase the effectiveness of using AI in CS education worldwide.

From a methodological perspective, the present study revealed that the AI effect on students' learning achievement in CS education might decrease over longer intervention periods. Therefore, researchers, pedagogues, and methodologists should rethink the AI intervention duration when designing their experiments or teaching approaches. In this context, Wang et al. (2023) highlighted that the effect of AI in education might be shaped by several confounding variables, including intervention duration, and more attention should be paid to these variables to accurately measure the AI effect and provide more generalizable results.

From a practical perspective, this present study revealed that AI can enhance students' learning achievement in CS education. It is therefore

important to raise awareness of the AI opportunities in education, thereby increasing its adoption and development in various CS disciplines. Additionally, the study revealed that developing a successful AI-powered educational technology is not an easy task because it should be treated as an ecosystem, where several confounding variables (education level, learning mode, intervention duration, geographical location, etc.) might impact the overall effectiveness of AI in education. This highlights the need for multidisciplinary collaboration (educators, computer scientists, data scientists, instructional designers, etc.) when developing AI systems in education generally and in CS education specifically. Finally, the findings of this study can contribute to achieving "quality education," which is the fourth sustainable development goal (SDG) of the United Nations that several organizations and universities are working towards.

5.2. Limitations and future directions

Despite the reliability of this study, it still has some limitations that should be acknowledged and further researched. First, the findings of this study are limited by the search keywords and databases. Therefore, future research can complement the present research by considering more databases and search keywords. Additionally, this study did not investigate the effect of AI roles (e. g., tutor, teacher assistant, peer, instructor, etc.) as a potential moderator of the AI effect

in CS education. Such information is crucial to understanding how AI should be integrated into CS education to increase learning experiences and outcomes. It also did not consider CS disciplines when investigating the AI effect. Future studies can therefore address this by taking, for instance, the CS classification of the Association for Computing Machinery (ACM). Finally, this study did not delve much into the technical aspects of the implemented AI technologies. Future studies should build on this to investigate how different AI techniques and algorithms might moderate the effect of AI in CS education.

Author's contributions

Ahmed Tlili: Methodology; Software; Writing (original draft); Writing (review and editing).

References

References with an asterisk (*) indicate studies included in the analysis.

AAAI. (2018). AAAI Launches "AI for K-12" Initiative in collaboration with the Computer Science Teachers Association (CSTA) and AI4All. https://m.aaai.org/Pressroom/Releases/release-18-0515.pdf

*Abbasi, S., & Kazi, H. (2014). Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention. *Asian Journal of Applied Science and Engineering*, *3* (2), 251-260.

*Abbasi, S., Kazi, H., & Hussaini, N. N. (2019). Effect of chatbot systems on student's learning outcomes. *Sylwan*, *163* (10), 49-63.

*Ahn, Y.-H., y Oh, E.-Y. (2024). Effects of the international training program for enhancing intelligent capabilities through blended learning on computational thinking, artificial intelligence competencies, and core competencies for the future society in graduate students. *Applied Sciences*, *14* (3). https://doi.org/10.3390/app14030991

*Al-Abdullatif, A. M., Al-Dokhny, A. A., & Drwish, A. M. (2023). Implementing the Bashayer chatbot in Saudi higher education: Measuring the influence on students' motivation and learning strategies. *Frontiers in psychology*, *14*, 1129070. https://doi.org/10.3389/fpsyg.2023.1129070

*Arsovic, B., & Stefanovic, N. (2020). E-learning based on the adaptive learning model: Case study in Serbia. *Sādhanā*, *45* (1), 266. https://doi.org/10.1007/s12046-020-01499-8

Barnes, T., Boyer, K., Hsiao, S. I. H., Le, N. T., & Sosnovsky, S. (2017). Preface for the special issue on AI-supported education in computer science. *International Journal of Artificial Intelligence in Education*, *27*, 1-4. https://doi.org/10.1007/s40593-016-0123-y

Borenstein, M. (2022). Comprehensive meta-analysis software. In M. Egger, J. P. T. Higgins, & G. D. Smith (Eds.), *Systematic reviews in health research: Meta-analysis in context* (pp. 535-548). BMJ Books. https://doi.org/10.1002/9781119099369.ch27

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, *1* (2), 97-111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). John Wiley & Sons.

Çağataylı, M., & Çelebi, E. (2022). Estimating academic success in higher education using big five personality traits, a machine learning approach. *Arab Journal Scientific Engineering*, 47, 1289-1298. https://doi.org/10.1007/s13369-021-05873-4

*Chang, Y. H., Chen, Y. Y., Chen, N. S., Lu, Y. T., & Fang, R. J. (2016). Yet another adaptive learning management system based on Felder and Silverman's learning styles and Mashup. *Eurasia Journal of Mathematics, Science and Technology Education*, *12* (5), 1273-1285. https://doi.org/10.12973/eurasia.2016.1512a

Crow, T., Luxton-Reilly, A., & Wuensche, B. (2018). Intelligent tutoring systems for programming education: A systematic review. In *Proceedings of the 20th Australasian Computing Education Conference* (pp. 53-62). Association for Computing Machinery.

*Cuong, N. D. H., Arch-Int, N., & Arch-Int, S. (2018). FUSE: A fuzzy-semantic framework for personalizing learning recommendations. *International Journal of Information Technology & Decision Making*, *17* (4), 1173-1201. https://doi.org/10.1142/S0219622018500220

Dever, D. A., Azevedo, R., Cloude, E. B., & Wiedbusch, M. (2020). The impact of autonomy and types of informational text presentations in game-based environments on learning: Converging multi-channel processes data and learning outcomes. *International Journal of Artificial Intelligence in Education*, *30* (4), 581-615. https://doi.org/10.1007/s40593-020-00215-1

*Eryılmaz, M., & Adabashi, A. (2020). Development of an intelligent tutoring system using Bayesian networks and fuzzy logic for a higher student academic performance. *Applied Sciences*, *10* (19), 6638. https://doi.org/10.3390/app10196638

*Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, *19* (1), 1-19. https://doi.org/10.1186/s41239-022-00362-6

*Farah, J. C., Spaenlehauer, B., Sharma, V., Rodríguez-Triana, M. J., Ingram, S., & Gillet, D. (2022). Impersonating chatbots in a code review exercise to teach software engineering best practices. In *2022 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1634-1642). IEEE.

*Fidan, M., & Gencel, N. (2022). Supporting the instructional videos with chatbot and peer feedback mechanisms in online learning: The effects on learning performance and intrinsic motivation. *Journal of Educational Computing Research*, *60* (7), 1716-1741. https://doi.org/10.1177/07356331221077901

Francisco, R. E., & de Oliveira, F. (2022). Intelligent tutoring system for computer science education and the use of artificial intelligence: A literature review. *CSEDU*, (1), 338-345.

García-Martínez, I., Fernández-Batanero, J., Fernández-Cerero, J., & León, S. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: Systematic review and meta-analysis. *Journal of New Approaches in Educational Research*, *12* (1), 171-197. https://doi.org/10.7821/naer.2023.1.1240

Gerdes, A., Heeren, B., Jeuring, J., & Van Binsbergen, L. T. (2017). Ask-Elle: An adaptable programming tutor for Haskell giving automated feedback. *International Journal of Artificial Intelligence in Education*, *27*, 65-100. https://doi.org/10.1007/s40593-015-0080-x

Goel, A. K., & Joyner, D. A. (2017). Using AI to teach AI: Lessons from an online AI class. *Ai Magazine*, *38* (2), 48-59. https://doi.org/10.1609/aimag.v38i2.2732

*Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2017). An educational system for learning search algorithms and automatically assessing student performance. *International Journal of Artificial Intelligence in Education*, *27*, 207-240. https://doi.org/10.1007/s40593-016-0116-x

*Hayashi, Y. (2019). Multiple pedagogical conversational agents to support learner-learner collaborative learning: Effects of splitting suggestion types. *Cognitive Systems Research*, *54*, 246-257. https://doi.org/10.1016/j.cogsys.2018.04.005

Hedges, L. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6* (2), 107-128. https://doi.org/10.3102/10769986006002107

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hwang, S. (2022). Examining the effects of artificial intelligence on elementary students' mathematics achievement: A meta-analysis. *Sustainability*, *14* (20), 1-20. https://doi.org/10.3390/su142013185

Jarota, M., (2023). Artificial intelligence in the work process. A reflection on the proposed European Union regulations on artificial intelligence from an occupational health and safety perspective. *Computer Law & Security Review*, *49*, 1-14. https://doi.org/10.1016/j.clsr.2023.105825

Kim, J., & Lee, SS. (2023). Are two heads better than one? The effect of student-AI collaboration on students' learning task performance. *TechTrends*, *67*, 365-375. https://doi.org/10.1007/s11528-022-00788-9

Kitchenham, B. A., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele University and Durham University.

*Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. *International journal of educational technology in higher education*, *18* (1), 1-28. https://doi.org/10.1186/s41239-021-00302-w

*Lai, Y. H., Chen, S. Y., Lai, C. F., Chang, Y. C., & Su, Y. S. (2021). Study on enhancing AIoT computational thinking skills by plot image-based VR. *Interactive Learning Environments*, *29* (3), 482-495. https://doi.org/10.1080/10494820.2019.1580750

*Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational technology research and development*, *70* (5), 1843-1865. https://doi.org/10.1007/s11423-022-10142-8

Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of General Internal Medicine*, *33* (8), 1260-1267. https://doi.org/10.1007/s11606-018-4425-7

Lin, R., Zhang, Q., Xi, L., & Chu, J. (2022). Exploring the effectiveness and moderators of artificial intelligence in the classroom: A meta-analysis. In J. Yang, D. Liu, Kinshuk, A. Tlili, M. Chang, E. Popescu, D. Burgos, & Z. Altınay (Eds.), *Resilience and future of smart learning. Proceedings of 2022 International Conference on Smart Learning Environments*. Springer. https://doi.org/10.1007/978-981-19-5967-7_7

Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments*, *10* (1), 41. https://doi.org/10.1186/s40561-023-00260-y

*Ma, Z. H., Hwang, W. Y., & Shih, T. K. (2020). Effects of a peer tutor recommender system (PTRS) with machine learning and automated assessment on vocational high school students' computer application operating skills. *Journal of Computers in Education*, *7*, 435-462. https://doi.org/10.1007/s40692-020-00162-9

Mihalic, S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth*, *4* (4), 83-105.

Merilampi, S., Sirkka, A., Leino, M., Koivisto, A. and Finn, E. (2014). Cognitive mobile games for memory impaired older adults. *Journal of Assistive Technologies*, *8* (4), 207-223. https://doi.org/10.1108/JAT-12-2013-0033

Nesbit, J. C., Adesope, O. O., Liu, Q., & Ma, W. (2014). How effective are intelligent tutoring systems in computer science education? In *The 14th IEEE International Conference on Advanced Learning Technologies (ICALT 2014). Advanced Technologies for Supporting Open Access to Formal and Informal Learning* (pp. 99-103). IEEE.

Nesbit, J., Liu, L., Liu, Q., and Adesope, O. O. (2015). Work in progress: Intelligent tutoring systems in computer science and software engineering education. In *2015 ASEE Annual Conference and Exposition Proceedings* (pp. 26.1754.1-26.1754.12). American Society for Engineering Education.

Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, *25*, 177-203. https://doi.org/10.1007/s40593-014-0028-6

Onesi-Ozigagun, O., Ololade, Y. J., Eyo-Udo, N. L., & Ogundipe, D. O. (2024). Revolutionizing education through AI: A comprehensive review of enhancing learning experiences. *International Journal of Applied Research in Social Sciences*, *6* (4), 589-607. https://doi.org/10.51594/ijarss.v6i4.1011

*Ortega-Ochoa, E., Quiroga, J., Arguedas, M., Daradoumis, T., & Marquès, J. M. (2024). The effectiveness of empathic chatbot feedback for developing computer competencies, motivation, self-regulation, and metacognitive reasoning in online higher education. *Internet of Things*, *25*, 101101. https://doi.org/10.1016/j.iot.2024.101101

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, M. A., Stewart, L. A., Thomas, J. Tricco, A. C., ..., y Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International journal of surgery*, *88*, 105906.

*Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2017). Assistance and feedback mechanism in an intelligent tutoring system for teaching conversion of natural language into logic. *International Journal of Artificial Intelligence in Education*, *27*, 475-514. https://doi.org/10.1007/s40593-017-0139-y

Pirker, J., Riffnaller-Schiefer, M., & Gütl, C. (2014). Motivational active learning: Engaging university students in computer science education. In *Proceedings of the 2014 Conference on Innovation y Technology in Computer Science Education* (pp. 297-302). Association for Computing Machinery.

Pisapia, J., Schlesinger, J., & Parks, A. (1993). *Learning technologies in the classroom: Review of the literature*. Metropolitan Educational Research Consortium.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86* (3), 638-641. https://doi.org/10.1037//0033-2909.86.3.638

Shaw, M. (1985). The nature of computer science. In M. Shaw (Ed.), *The Carnegie-Mellon curriculum for undergraduate computer science*. Springer.

Silva, D. B., Aguiar, R. d. L., Dvconlo, D. S., & Silla, C. N. (2019). Recent studies about teaching algorithms (CS1) and data structures (CS2) for computer science students. In *2019 IEEE Frontiers in Education Conference* (FIE) (pp. 1-8). IEEE.

Smith, M. (2016, 30 January). Computer science for all [blog]. *The White House - President Barack Obama*. https://www.whitehouse.gov/blog/2016/01/30/computer-science-all

*Song, D., & Kim, D. (2021). Effects of self-regulation scaffolding on online participation and learning outcomes. *Journal of Research on Technology in Education*, *53* (3), 249-263. https://doi.org/10.1080/15391523.2020.1767525

*Stamper, J., Eagle, M., Barnes, T., & Croy, M. (2013). Experimental evaluation of automatic hint generation for a logic tutor. *International Journal of Artificial Intelligence in Education*, *22* (1-2), 3-17. https://doi.org/10.1007/978-3-642-21869-9_45

*Su, J. M. (2020). A rule-based self-regulated learning assistance scheme to facilitate personalized learning with adaptive scaffoldings: A case study for learning computer software. *Computer Applications in Engineering Education*, *28* (3), 536-555. https://doi.org/10.1002/cae.22222

*Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, *21*, 14. https://doi.org/10.1186/s41239-024-00446-5

Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, *94*, 252-275. https://doi.org/10.1016/j.compedu.2015.11.008

Tangi, L., van Noordt, C., Combetto, M., Gattwinkel D., & Pignatelli F. (2022). *AI watch. European landscape on the use of artificial intelligence by the public sector*. Publications Office of the European Union. https://doi.org/10.2760/39336

*Tegos, S., & Demetriadis, S. (2017). Conversational agents improve peer learning through building on prior knowledge. *Journal of Educational Technology & Society*, *20* (1), 99-111.

*Tegos, S., Demetriadis, S., & Karakostas, A. (2015). Promoting academically productive talk with conversational agent interventions in collaborative learning settings. *Computers & Education*, *87*, 309-325. https://doi.org/10.1016/j.compedu.2015.07.014

*Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: A comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, *11* (4), 417-440. https://doi.org/10.1007/s11412-016-9246-2

Tlili, A., Essalmi, F., & Jemni, M. (2015). An educational game for teaching computer architecture: Evaluation using learning analytics. In *2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA)* (pp. 1-6). IEEE.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10* (1), 15. https://doi.org/10.1186/s40561-023-00237-x

Touretzky, D., Gardner-McCune, C., & Seehorn, D. (2023). Machine learning and the five big ideas in AI. *International Journal of Artificial Intelligence in Education*, *33* (2), 233-266. https://doi.org/10.1007/s40593-022-00314-1

Verdú, E., Regueras, L. M., Gal, E., De Castro, J. P., Verdú, M. J., & Kohen-Vacs, D. (2017). Integration of an intelligent tutoring system in a course of computer network design. *Educational Technology Research and Development*, *65*, 653-677. https://doi.org/10.1007/s11423-016-9503-0

Waraich, A. (2004). Using narrative as a motivating device to teach binary arithmetic and logic gates. *ACM SIGCSE Bulletin*, *36* (3), 97-101. https://doi.org/10.1145/1026487.1008024

Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and information technologies*, *28* (7), 9113-9148. https://doi.org/10.1007/s10639-022-11555-x

*Winkler, R., Hobert, S., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020, April). Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). Association for Computing Machinery.

Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55* (1), 10-33. https://doi.org/10.1111/bjet.13334

*Yin, J., Goh, T. T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, *59* (1), 154-177. https://doi.org/10.1177/0735633120952067

Zhang, X., Tlili, A., Shubeck, K., Hu, X., Huang, R., & Zhu, L. (2021). Teachers' adoption of an open and interactive e-book for teaching K-12 students artificial intelligence: A mixed methods inquiry. *Smart Learning Environments*, *8*, 1-20. https://doi.org/10.1186/s40561-021-00176-5

Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2023). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, *31* (9), 5650-5664. https://doi.org/10.1080/10494820.2021.2015693

Author's biography

Ahmed Tlili. Associate professor of Educational Technology at Beijing Normal University, China, and adjunct associate professor at An-Najah University, Palestine. He is the co-director of the OER Lab at the Smart Learning Institute of Beijing Normal University (SLIBNU), China. He serves as the editor of Springer Series *Future Education and Learning Spaces*, and the deputy-editor-in-chief of *Smart Learning Environments* journal. He is also the associate editor of *IEEE Bulletin of the Technical Committee on Learning Technology*, and the journal of *e-Learning and Knowledge Society*. Professor Tlili is also an expert at the Arab League Educational, Cultural and Scientific Organization (ALECSO). He has edited several special issues in several journals. He has also published several books, as well as academic papers in international referred journals and conferences. He has been awarded the Martin Wolpers 2021 Prize by the Research Institute for Innovation and Technology in Education (UNIR iTED) in recognition of excellence in research, education and significant impact on society. He also has been awarded the IEEE TCLT Early Career Researcher Award in Learning Technologies for 2020.

https://orcid.org/0000-0003-1449-7751

Teaching pro-environmental rules: Implications for environmental education and climate change

La formación en reglas proambientales: implicaciones para la educación ambiental y el cambio climático

Leidy-Vanessa DÍAZ-BELTRÁN[1]. Teacher. Secretariat of Education, Colombia (*vanessadiazb@gmail.com*).
Pablo PÁRAMO, PhD. Professor. Universidad Pedagógica Nacional, Colombia (*pdeparamo@gmail.com*).

Abstract:

The concern regarding the gap between *words* and *actions* has spread through different contexts, in particular those relating to environmental education. Despite repeated warnings about the damaging impact of human behaviour on the environment and the effect of actions undertaken to teach people to protect the environment (pro-environmental behaviours), there have been few advances in relation to favourable responses from people's *actions*. The alternatives proposed in environmental education to reduce this gap include research relating to teaching pro-environmental rules and its impact on human behaviour. Rule learning poses challenges for the comprehension of human behaviour, especially for environmental education in general and climate change in particular as well as being proposed as an epistemological and theoretical approach to human behaviour insofar as it differs from the established approach in physiological, medical, attitudinal, cognitive, or motivational mediation models that explain behaviour on the basis of factors that are internal to the individual. This article discusses the origins, genetic or environmental [Chomsky (1959) *vs.* Skinner (1957, 1981)], of verbal behaviour with regards to its origin and development in ontogenesis. Afterwards, it considers the notion of *rule-governed behaviour* with the definitions and different taxonomies that have been proposed regarding rules, as well as the variables associated with (in)sensitivity to following them and the possible alternative approaches for each of the (in)sensitivity factors identified. Finally, it analyses the role of education in pro-environmental rules for regulating people's own behaviour and in the design of macro- and

metacontingencies to incentivise environmental protection in large human groups.

Keywords: pro-environmental behaviour, universal grammar, verbal regulation, rules-governed behaviour, environmental education.

Resumen:

La preocupación causada por la brecha existente entre el *decir* y el *hacer* permea diferentes contextos, en particular aquellos relacionados con la educación ambiental. A pesar de las reiteradas advertencias sobre las implicaciones del comportamiento humano en el detrimento ambiental y de las acciones emprendidas para la formación en pro de la protección del medio ambiente (comportamientos proambientales), pocos son los avances que se han logrado con relación a las respuestas favorables desde el *hacer* de las personas. Entre las alternativas planteadas en la educación ambiental para reducir esta brecha se encuentra la búsqueda relacionada con la formación en reglas proambientales y su implicación en el comportamiento humano. El aprendizaje por reglas no solo se propone como aproximación epistemológica y teórica del comportamiento humano. Además, en la medida en que se aleja de la visión establecida en modelos mediacionales fisiológicos, médicos, actitudinales, cognitivos o motivacionales (que explican el comportamiento a partir de factores internos del individuo), plantea retos para la comprensión del comportamiento humano; en especial para la educación ambiental en general y para el cambio climático en particular. El presente artículo discute los orígenes, genéticos o ambientales [Chomsky (1959) *vs.* Skinner (1957, 1981)], de la conducta verbal en cuanto a su nacimiento y desarrollo en la ontogénesis. A continuación, aborda la noción *conducta gobernada por reglas* con las definiciones y diferentes taxonomías que se han propuesto sobre las reglas, así como las variables asociadas a la (in)sensibilidad para seguirlas y los posibles abordajes alternativos a cada uno de los factores de la (in)sensibilidad identificados. Por último, analiza el papel que juega la educación en reglas proambientales en la regulación del propio comportamiento, así como en el diseño de macro- y metacontingencias para incentivar la protección del medio ambiente en grupos humanos amplios.

Palabras clave: comportamiento proambiental, gramática universal, regulación verbal, conducta gobernada por reglas, educación ambiental.

1. The origin of language

Although different learning mechanisms for explaining behaviour have been explored, such as learning by association (Pavlov, in Gutiérrez 2005), action by direct consequences in the environment, or operant conditioning (Skinner, 1938, 1953), imitation (Bandura & Walters, 1959), and cognition (Broadbent, 1958; Neisser, 1967; Piaget, 1978), mechanisms that are largely shared with non-human animal species, the principal mechanism

by which people learn is verbal regulation, which occurs through the use of language.

Within evolutionary theory there are various hypotheses regarding how communication systems might have evolved in different animal species, many of which are contradictory; however, all of these hypotheses agree that language is a specific capacity of the human being that exceeds the ones perceived in non-human animal behaviour and is the principal mechanism in the creation of culture.

Attributing verbal behaviour (not communication) exclusively to the human species recognises the difficulty of accurately deciphering language's origin owing to the lack of records and to the scarcity of evidence (it does not leave a fossil record), with there being only a few products such as writing. Archaeological proofs only date back to the origin of writing, some five thousand years ago, a period in which the critical genetic changes in *Homo sapiens* had already occurred and the rules of speech had already been established uniformly in all of the world's societies. However, there are a few patterns in speech that can be cited as products of evolution. One of these traces is the rhythm of turn-taking in conversations; when the conversational intervals of speakers of ten languages from all over the world were measured, it was found that they all avoided overlapping (but not interruption), and it was observed that the duration of the time between conversation and conversation was almost the same (Sauter et al., 2010).

Another trace of early linguistic evolution that has recently been documented is found in non-verbal vocalisations, exclamations that are probably older than language. It has, for example, been observed that vocalisations that communicate negative emotions (anger, disgust, fear, and sadness) are the same for native speakers of English in Europe and for speakers of the Himba language, which is only spoken in a few culturally isolated villages in northern Namibia. In contrast, non-verbal vocalisations that communicate positive emotions (achievements, fun, sensual pleasure, and relief) do not match in the same way, although the reasons for this difference are unknown (Sauter et al., 2010).

Consequently, it is clear that from an early-stage human beings started to display sensitivity to social interactions or contingencies2 in the development of language. In this way, it is important to underline the role of operant conditioning (Skinner, 1986), which suggests that the conditions for the emergence of verbal behaviour arose when the capacity to exercise control over the phonological apparatus was selected for by evolution, not as a result of natural selection but of the verbal social setting and its variations.

From this perspective, although the existence of verbal behaviour is initially conditioned on the genetic changes that the human species undergoes in most of its vocal organs, starting with a low larynx that makes vocalisation possible, a feature that is absent in our primate

cousins and in children aged under two years (Lieberman, 1998), language emerges with the development and operant control of the phonological articulation apparatus. This difference, in physiological changes in humans stands out even more for "an increase in general learning abilities or intelligence" (Álvarez-González, 2010, p. 5), and for humans' condition as social beings. All this marks an advance in verbal behaviour that goes beyond genetics and leads to it being controlled by the consequences of the social environment linked to action and imitation.

Humans have had evolutionary success thanks to the social skills that they have developed; the crucial difference between humans and other animal species, including our closest genetic relatives, chimpanzees, is the capacity to collaborate to achieve shared goals and intentions (Tomasello, 2016).

This evolutionary stage of the human species emphasises the role that social reinforcement contingencies fulfil on human behaviour on the basis of modelling, that is, the exhibition of the behaviour of one individual in front of another. In this process, the way people act is, on the one hand, linked to the interest in achieving the same as the person who modelled the behaviour; on the other, to the possibility that the behaviour can be generated or inhibited through a gesture without the need for the listener to experience the result of the behaviour directly. Within this explanation, the prior experience of giving instruction

also stands out so that the listener, acting on these instructions, receives reinforcing contingencies that will at the same time be reciprocal for the speaker. In this way, the evolution of language is not related to organs or to knowledge, but to the cultural milieu.

In contrast with this approach, Chomsky, from biological and linguistic anthropology, delegates language principally to innate linguistic structures, the result of biological evolution and natural selection, which he synthesises in a set of norms on which he bases the existence of a deep or universal grammar in the brain. According to him, human beings possess an innate language acquisition module (Chomsky, 1959). Chomsky argues that a language with all of its added grammatical rules is too complex for a child to be able to learn its grammatical structure in such a short period (children start speaking between 12 and 18 months in age and are very soon able to construct their own phrases and correct themselves grammatically). The phoneme that a child hears from its mother in its first years of life takes on a different meaning and triggers markedly different responses to those that the same phoneme might trigger in a young chimpanzee, something that can be attributed to the linguistic structure that precedes language in children.

Chomsky not only criticises Skinner but also Wittgenstein and Quinne, who stress social influence on language, as this seems superficial and trivial to him

because it is impossible to carry out a scientific analysis of the social influence on the development of language given the many variables that are at play when we use it (GramaPsico, 2023). Chomsky also supports his point of view by elaborating a series of norms that, as he claimed, albeit lacking further evidence (and somewhat unclearly), are spontaneously followed in the developing brain, minimising the role of the environment and learning in language acquisition.

The criticisms Chomsky made were refuted by MacCorquodale (1969), who established that they were irrelevant owing to Chomsky's lack of understanding of Skinner's proposed approach. According to MacCorquodale, Chomsky's comments on Skinner basically centred on the style of the work and on references to Skinner's work with rats. Thus, he left aside valuable arguments Skinner made when he removed the individual's dualist condition, according to which the individual exercises autonomy through thought. Skinner argued that thought is not the cause of behaviour, including verbal behaviour, but that it is also the result of operant conditioning, in particular through social contingencies or consequences.

The speaker is not an autonomous agent, Skinner argues in his book *Verbal behavior*. Verbal behaviour is regulated by the multitude of effects that can be obtained within the variations of the social setting, thus differing from the *tabula rasa* vision of Watson and Hull's cognitive one, on which

Chomsky apparently based his critique. Skinner considers verbal behaviour to be the result of what is provided by a regulating variable, social contingency, that is represented by listening and by supplementary variables such as education through the use of books, music, educational games, and the positive reinforcement that parents provide when children use language correctly in family conversation.

After many debates about the innate or cultural character of the origin of language, it is currently accepted that there is an instinctive component in language learning, a predisposition, as Pinker (2007) argues, but also that there is a sensitive period in the child's development where this learning can happen quickly, and it also accepted that that the existence of a special module for grammar in the brain is not an unavoidable necessity for explaining it. As for many other behaviours, there is a "preparation for learning", or pre-adaptations, as shown in various studies with animals collected in the works by Seligman and Hager (1972) and Marler and Terrace (1984) and the observations of the sociobiologists (Wilson, 2000; Alcock, 2001), but with few limitations for the development, particularly of language in humans.

One useful concept here is that of closed and open instincts and programs (Midgley, 1978). According to the author, closed instincts are behavioural patterns that are fixed genetically in every detail, such as the honey dance of bees,

the song of some birds, and the pattern of nest building of weaverbirds. Open instincts, on the other hand, are behavioural programs that have a void. Parts of the behavioural pattern are innate, determined, but others are completed in the programme with experience. Experience must *imprint* it. This is the simplest type of open instinct. However, even the simplest animals carry out activities in which the space left for experience has to be much broader, but the general objective is still innately determined. This appears to be the case with language and with many other behaviours in both humans and other non-human animal species.

As Wilson (2012) states, there seems to be a genetic rule biased towards the subject, verb, object word order integrated into our biological structure, although its end products in grammar are very flexible and are learned. The multitude of pathways in the evolution of basic syntax suggests that there are few if any genetic rules that guide the learning of language by individual human beings. So, as Wilson states, it seems that Chomsky is partly right, but Skinner is more so when he emphasises learning.

Why is there so little emphasis on the innate component of language? Because the rapidly changing environment of speech does not provide a stable environment for it to have been affected by natural selection. Language varies too quickly over generations and from one culture to another for it to have been the results of evolution by selection. Consequently,

there are few reasons to expect that the arbitrary properties of language, including the abstract syntactical principles of the phrases and the mark of genes, have been integrated in a special "language module" in the brain through the process of natural selection, according to Wilson (2012).

As a result, scholars of the origin of language from the innatist structural perspective have qualified their approaches when determining the malleability of this structure on the basis of the influence of learning and its critical periods, thus coming closer to the assumptions of Skinner. So, the dispute that emerged between these two authors in 1959 relating to language acquisition transcends the debate about whether language is learned or innate, since as MacCorquodale (1969) argues, regardless of the innatist or operant (structural or functional) perspective, language is a complex matter, full of phonemes and dynamic relations with the environment that cannot be limited to a grammatical standard, which leads us to analyse how language develops in ontogenesis, and to perform a functional analysis from the influence that it exercises on the social environment.

2. Acquisition of verbal behaviour in ontogenesis

After considering the debate about the origin of verbal behaviour, our attention now turns to functional analysis of verbal behaviour (Skinner 1981, 1986), principally considering the sensitivity to the influence of consequences on behaviour

(operant conditioning) and social reinforcement, as the principal conditioner of the verbal response. In this process the dual role that each person plays as a speaker and a listener in social reinforcement is significant. It is notable that, while behaviour after being generated or avoided by social reinforcement can occur in the absence of the listener, it is more likely to be maintained in the listener's presence owing to the reinforcement that the listener can issue. The dynamic characteristics that Skinner identified relating to verbal behaviour, as well as emphasising the role of speaker and listener, insist on both the inclusion of other languages (signs, gestures, writing) that can essentially be considered verbal because the speaker can shape the behaviour of the listener through them, and on the distinction of the operant behaviour considered and the response, prediction, or control that can be achieved in the listener's behaviour. Unlike Vygotsky (1986) and Luria (1961), Skinner does not distinguish between thought and language as for him thought is concealed verbal behaviour.

3. Verbal regulation

What role does language play in the behaviour of others and on one's own? According to Catania (2003), people's susceptibility to their behaviour being verbally regulated can be shaped by the power that words acquire in verbal behaviour. Such power depends on their function of influencing others' behaviour, both verbal and non-verbal, by means of instructions, suggestions,

descriptions, advice, indications of consequences or contingencies, laws, social values, social norms, and self-generated verbal rules.

This is how social contingencies or social consequences in verbal regulation not only involve changing the behaviour of the person who listens or speaks but also the fit of the correspondence between words and action. Contingencies that affect verbal behaviour are often different to those that prevail in non-verbal behaviour. For example, we can be celebrated socially for speaking in favour of the environment and censured if we do not do so, which maintains favourable opinions in support of the environment. However, something different happens with non-verbal behaviour relating to protecting the environment. There is often no clear response based on whether we do something in favour of the environment or stop doing it (reducing the cost of the waste collection service or imposing fines, for example). Accordingly, the consequences or contingencies for both types of behaviour, verbal and non-verbal, are different and so a discrepancy between what we say and what we do emerges.

From this perspective, educational actions, especially those relating to education about pro-environmental behaviours, have generally been oriented towards teaching values and attitudes (understood as a person's knowledge and opinions about something) with the aim of influencing the person's behaviour within a dualist perspective where values, knowledge, and attitudes are the

rep

cause of the behaviour, its effect. Even so, results have shown that people do not always behave in accordance with their knowledge, their values, and their attitudes (Chawla & Derr, 2012). Hence, direct experience with the natural environment has been proposed as a more effective educational strategy (Collado & Corraliza, 2011), which nonetheless requires a clearer explanation of why it works.

These approaches seem insufficient when a central element in the change in behaviour is disregarded: the consequences. According to Páramo (2017), an epistemological alternative to tackling this dissonance between words and action regarding environmental education is to build a bridge between behaviour and consequences, for which the author proposes teaching rules. These, according to Ribes (2000) and Glenn (2003), are conceived as verbal descriptions that establish relations of dependency between situations where a behaviour is required, the behaviour itself, and the result it produces in the environment. The rule generally indicates by means of a verbal statement what an individual or group is expected to do in a particular situation and what would happen when doing it or not doing, even if experiencing its direct consequence is not necessary for following it, which is as a result of previous successful experiences when obeying them, or of the observation of what other do or cease to do (Gómez et al., 2006; Páramo, 2010, 2013, 2017; Baum, 2017; Kissi et al., 2017; Páramo & Burbano,

2019; Pietras, 2022; Páramo et al., in press).

Furthermore, it is surprising that people follow rules when the consequences of human actions are usually distant: there are nine months from conception to birth; the values and general rules that we learn as children at home mainly have an effect as adults. What would happen if we had to learn the traffic rules every morning, by means of trial and punishment? According to Baum (2017), even though humans preserve few memories of learning rules as children and their role in these first stages of life in their family and cultural setting is limited to listening and obeying, with the passage of time and after the age of approximately four, the effectiveness of verbal statements can be perceived, not only in following them but also in generalising them to other situations, transcending the passive figure of the listener, with the person becoming a reference, instructor, follower, and creator of rules.

In other words, a behaviour that is generally favoured by another person and that is subject to reinforcement or punishment (that is to say, that is verbally regulated) can over time easily fall into natural contingencies and finally become a behaviour that is implicitly moulded. This is the case with how we catch a ball or learn to ride a bicycle, where despite doing the action, we cannot describe in any detail why we do it one way and not another (Baum, 2017). Primary teachers often prefer not giving

rules or instructions directly but instead showing the evidence or consequences of following it and letting children make their own decisions. They even design the rules jointly with their students, as an experience that encourages them to build the rules in their words, revealing them on the basis of their consequences, and putting them into practice so that they are socially rewarded or receive direct contingencies.

How does verbal regulation act on the behaviour of the individual? While it is true that people learn to follow rules by directions from others or by imitation, it is apparent that they are also capable of verbally generating a rule based on their own experiences, something that leads to them being able to regulate their own behaviour (Peláez & Moreno, 1999). In contrast with what has been proposed from the cognitive outlook, in which reference is made to self-regulation as a characteristic or metacognitive skill of the individual to control his or her own behaviour and thoughts and plan his or her actions (De la Fuente et al. 2019), self-regulation from learning verbal rules emerges in the moment when the individual not only learns to follow rules whether through instructions, reference to other people, or observation, but also on the basis of the ones that the individual develops from prior experiences and the results of it.

So, in general rules form the principal motor of what is usually called *self-regulation*; in other words, insofar as individuals learn rules, these regu-

late their behaviour without direct external control necessarily mattering. This is how we move away from the approaches relating to self-regulation from the dualist perspective of Zimmerman (2008), De la Fuente (2017), and Panadero (2017). They place the cause of regulation on an internal mental level (something that gives rise to a circular or tautological definition, as the prefix *self-* involves an internal agent that regulates an individual's action) or on an instigator element external to the individual that is required for this to occur or to motivate the regulation of a given behaviour.

Instead, we consider it to be vitally important to note that human beings achieve true self-regulation through verbal regulation. The vision of rule-governed behaviour does not entail an internal effort as motivator of the action but instead positions the explanation in what direct or referred experience gives individuals to guide their behaviour. This makes speaking of self-regulation unnecessary unless it is understood to be the individual guiding its behaviour by the rules they have learnt.

4. Taxonomies of rules

With the aim of developing the exposition of pro-environmental rules for people's environmental education, the present epistemological approach first provides analyses of how rules have been classified according to their theoretical or practical purposes. As a result, the following distinction between taxonomies

of rules is established according to: how the society is organised; how the rules are expressed; their social function; the formative reference point; or the type of contingency associated with following them.

To start, rules are distinguished according to how the society has been organised: legal, social, or moral. The first type refers to regulations that are established by the authorities of a society, based on codes, constitutions, laws, and other provisions; changes in them are easy to perceive and disobeying them results in legal punishments. The second type, social norms, corresponds to social customs (culture) under which a social group generates recognition or reproach in relation to the behaviour of a person. And finally, moral rules, through which a person regulates his or her behaviour; these are conditioned by the person's beliefs of what is considered good or bad (feelings of guilt) (Bermúdez, 2021). It may however be the case that some are interlinked. Someone who cuts down a tree can incur legal sanctions because this is classified by law as an environmental offence and he must pay a fine. At the same time, he can receive a social sanction by being judged by the community, and he might also experience remorse because he has acted against his morals.

Moreover, rules are recognised according to how they are expressed: tacitly or explicitly. The first type is communicated symbolically, which means they can be learnt on the basis of the education received with regards to the symbol with which they are transmitted, and so will be followed by observation of the symbol. These are immersed in and limited by each culture by means of language, whether of courtesy or for conventions relating to the moral, the religious, the environmental, etc. For example, a sign prohibiting throwing waste into water sources. The second type are explicitly or formally expressed, which means they include a physical instigator in the environment. They are generally related to the law or to rights, and so they are embodied through police codes or codes for coexistence (Páramo, 2010). This can be illustrated by the following statement: "If you leave rubbish here, you will be punished with a fine of the minimum wage"; here the consequence is stated precisely.

Rules can also be designed and perceived according to their social function, and so they are generally aimed at coexistence and care for the environment (Páramo, 2013; Páramo et al., in press). These include ones that relate to hygiene, security, prohibition, solidarity, respect, prevention, mobility, and concern for the individual and collective health as presented in police codes or in urban signage.

The distinction between rules based on their formative reference point should also be noted. Cialdini et al. (1990) mention three types of norms (rules in our case) that can be included in this distinction. First,

descriptive norms, which are closely linked to observing what most people generally do; the importance of social influence, which makes them effective and adaptive, is recognised on the basis of them. Their principal motivation is coexistence, and they are represented by social norms and environmental beliefs. They can, for example, be perceived when most people in public places put their rubbish in the containers provided for this purpose, avoiding contamination and promoting healthy coexistence.

The second are *prescriptive* norms, of which their formative social influence on what is expected of the other stands out, that is to say, what is desirable. These are generally inspired by the feelings or beliefs of those considered correct or incorrect (that is to say, what corresponds to the moral obligation) and motivated by anticipated prizes and punishments. They can be perceived in the actions undertaken in recycling campaigns, where an instruction is given and the reason for following it is explained. For example, the statement "Dispose of plastic bottles here and avoid contaminating our planet" accompanied by a happy face or a thumbs up.

Lastly, the third are *personal prescriptive* norms, which derive from social moral judgement or (dis)approval of what is regarded as correct or incorrect. Here people judge their own behaviour according to their beliefs. So, it is natural for people to feel morally obliged to care for the environment if this starts

from the beliefs with which they were educated.

Finally, and of great relevance for the purpose of the present article, rules can also be defined based on the contingency (consequences of following them) (Hayes et al., 2001; Gómez et al., 2006; Kissi et al., 2017; Páramo & Burbano, 2019; Harte et al., 2020; Pietras, 2022; Páramo et al., in press). According to these authors, the behaviour that the rule establishes is followed because of the type of contingency linked to it: social approval, the natural consequence with which the rule is linked, or the increased benefit of following it. Accordingly, there are three types of rule-governed behaviours: the first type relates to rule-following driven by social reinforcement, which has been called *pliance* (obedience / consent or compliance). Here people's actions are mediated by social recognition or criticism, or by the consequences for the speaker of following the rules. These can be found in statements such as "If you plant trees, you will receive a voucher" or "Get the certificate for the best environmental education project". Here the consequence is furnished by another person, the rule is conceived as a social regulation mechanism.

The second rule-following behaviour is associated with obeying rules because of natural or intrinsic benefit, which obeys the way in which the world is organised and has been called *tracking*. In this case, rule-following behaviour occurs because of the reinforcement of the correspondence be-

tween the rule and the natural consequences of following it. For example, assembling a table by following the instructions that come with it; following instructions to solve a mathematical problem is reinforced by the logical result that derives from following the procedure. In the same way, rule-following because of natural contingencies is apparent when saying "Disposing of waste in the right place prevents blocked drains and helps conserve water sources" and in rainy periods an improvement in prevention of the floods that generally occur is noted. Here, rule-following relates to what Peláez and Moreno (1999) call the *accuracy of a rule*, in other words, when the contingency that is obtained matches what was said would be obtained from following the rule. This correspondence is what increases the possibility of it being followed and what results in it being called an accurate rule.

And the third rule-following behaviour is associated with the reinforcement given, from the rules with increased motivational effect (Kissi et al. 2017), known as *augmenting* (behaviour shaped by changes such as prizes and punishments) artificially augmented. Here the value of the reinforcer in the previous two types is augmented from motivative augments and formative augments. The first type refers to rules that have an effect both on the value of stimuli and on the alteration of behaviour. This is how the possibility of following a certain behaviour is increased or reduced from

the rule by magnifying the consequence of this, for example, "Don't use your vehicle, you are burning our planet". Meanwhile, from the second type, new reinforcers or punishments emerge; although these are less common than the previous ones, they have a notable role in the control of behaviour regulated by rules. This can be illustrated by the following advertising slogans: "Our rivers are not your waste heap, prison await you" or "Redeem your voucher, recycling pays well".

As Bermúdez (2021) notes, laws, which are nothing more than verbal rules, undoubtedly become the principal mechanism for regulation of a society. This is of course without ignoring the criticism that derives from the totalitarianism from which legal norms sometimes arise and following them when people move away from their principles and submit themselves to the power exercised by different systems (religious, political, etc.), prioritising obedience and authoritarianism, or ideology. This makes visible the existing imbalance between norms, sometimes emphasising the social and excluding the moral, or emphasising the moral aspect but leaving aside the legal, which can be seen in everyday matters of societies as it is a behaviour that is socially punished but even so, is not susceptible to a legal sanction. Social and moral rules can strengthen or challenge what is established from the legal sphere. This variability or contradiction between laws can in some way contribute to pliance with them and their effectiveness, both

in the individual sphere and in group behaviours, something that will be addressed in the following section.

5. (In)sensitivity to following rules: how can this be tackled?

Having considered the origins of verbal behaviour, its influence on human behaviour, and the way attempts are made through verbal regulation to educate and change people's behaviour by enunciating verbal rules, we now come to the question of why despite the benefits of rule-following, whether to achieve outcomes that favour people or prevent undesirable situations for them, some people do not follow them; in other words, are insensitive to learning them, something called *insensitivity to rule following* (Kissi et al., 2017). "This insensitivity refers to the relative absence of control through consequences, because the behaviour is assumed to be sensitive to the contingencies that shaped the rule-following behaviour" (Cerutti, 1989, p. 260); or, according to Gómez et al. (2006), "to the lack of fit to the direct contingencies, as the behaviour would be under the control of words and would be sensitive to these" (p. 59).

Skinner (1981, 1986) considers this aspect from two perspectives. He contemplates verbal behaviour as a dependent variable, where, on the one hand, some verbal stimuli (which, according to his theory, would be seen as discriminative stimuli3) could be inadequate, unclear, or poorly presented or, on the other hand, could present an absence of reinforcement or be delayed. Although the behaviour can

be maintained despite the absence of a direct or social reinforcement, over time this tends to weaken and even disappear, as the author states.

A more in-depth examination of these possible situations is presented below in which (in)sensitivity to contingencies can be displayed from in two categories, according to Luciano (as cited in Gómez et al., 2006):

1. *Discrimination between contingencies and what is described in the rule*, a category that entails confusion between what the contingency establishes and what the rule says.

2. *Competition between direct consequences of doing something and the social consequences of rule following*, a category that entails a difference between the contingencies of personal benefit received.

Enrichment of this perspective is pursued, adding new considerations to these categories and incorporating some aspects identified by authors such as Peláez and Moreno (1999), Gifford (2011), Bermúdez (2021), Pinker (2021), and Pietras (2022), in turn proposing educational options for resolving it.

On the one hand, the first category *(discrimination between contingencies and what is described in the rule)* groups factors such as:

a) Variability / inconsistency between the rule and its consequence: this can, for

example, be seen in intergenerational change, when older people expect young people to act how they did, even when the contingencies are no longer the same. So, it becomes necessary to make changes to contingencies for new generations for behaviours that do not fit the rules.

b) Ambiguous content of the rule: the rule is not clear, there is a discrepancy between what is established verbally and the consequences that occur. To resolve this, making explicit the relationship between the situation, the expected behaviour, and the consequence is proposed.

c) Desensitisation: while a direct relationship is perceived between the flexibility or the rigour of the rule and the disorder or self-control that defines societies, this will be unable to become the fundamental factor for arguing how people should behave. The contingency for following the rule loses value if there is overexposure to it, and so it is necessary to replace the sanctions for not following the rules with social recognition and offer benefits of obeying them.

d) Personal history: some people / students are more sensitive to rule following if they experience the contingency, some are more sensitive to observing experience of the contingency, and others are more sensitive to the instruction given in order to receive the contingency. Therefore, a type of rule can function for some individuals

but not for others. From this perspective, variability in people's behaviour should be generated before changing the contingency.

e) Immediate or delayed consequences after verbalising the rule: immediacy in receiving the consequence is a determining factor in the sensitivity to following rules. This can, for example, be evinced in relation to the implications of climate change, which might be experienced in the very long term, resulting in a low likelihood of the rule being followed, in contrast with what happens when the consequence is immediate, which guarantees the rule will be followed. One way of approaching this situation might be to increase the value of the signs that announce benefits or positive or negative consequences by making the consequence closer in time based on signs that announce its closeness. For example, in the case of climate change, showing the effects of floods or extreme droughts.

f) Consequences of execution of the rule: this aspect is related to the reinforcement or punishment value of the contingency. In some cases, the contingency does not exist or is poor, which results in the rule not being followed. It can also be perceived through doubt or lack of certainty in relation to the efficacy of the actions undertaken in relation to climate change. So, the question of how effective/true (possible) the contingency arises, which is called the *degree of fit between the*

rules and the contingencies. This aspect relates to the likelihood of receiving the contingency. Promoting behaviour change by making people see that following rules effectively produces consequences, whether positive or negative, is proposed. It then becomes necessary to ensure that the consequences are received and that their value is sufficiently valuable or punitive for the individual.

g) Lack of experience of following rules: for people to be more sensitive, teaching rules and following them from an early age is proposed in different educational scenarios: the home or the school.

h) Lack of dissemination of laws or ignorance of the problems from which they emerge: one factor that has a major impact on pliance with rules is knowledge of them or recognition of their impact in the face of a current problem; not knowing of their existence means society is less likely to comply with them. To counteract the shortcoming, expanding the channels of communication is proposed, thus guaranteeing cover of the different social actors.

i) Differences in culture or rule-following traditions: the rule is decontextualised or goes against the culture. One way to approach this dissonance is to take into account the moral beliefs of the society when designing of rules, which will without any doubt facilitate their fulfilment.

j) Irrationality in the face of the systematic (the rules of reason and science): there is no appetite for or interest in what involves bias. One possible approach is to promote rationality from early childhood by designing rules that fit logic and scientific evidence. Agreeing on rules that realign incentives towards the truth. Valuing the norm of rationality itself/approval or disapproval of rational or irrational habits.

k) Contradictory norms: distortion or lack of scientific, journalistic, or political consensus reduces belief in certain factors. For this, the proposal is to disseminate messages of consensus, which are a method for overcoming the debilitating effect. This factor can also be perceived in contradiction between the same norms, in other words, social, legal, and moral norms contradicting one another. Here it is necessary to work against obedience and shape individuals' critical thinking so that they can determine priorities and decide what is relevant when facing the paradoxes that are presented within the rules.

l) The contingency for following the rule is artificial: this aspect alludes to a lack of correspondence between following a rule and receiving a natural contingency or one that is inherent to how society is organised. If an individual gets a hamburger as an incentive for reducing electricity use, the difficulty of establishing a logical and natural connection between pliance with the

rule and the contingencies of doing so is clear. For this reason, the importance of establishing a coherent relationship between the rule and the contingency is identified.

m) Lack of credibility of the source from which the rule or norm emerges: Although there is no correspondence with the contingencies from following the rule or they are unknown, people are more inclined to follow a rule the more social recognition for following it is perceived or if there is a generalised favourable image of the situation that the rule imposes, which ensures that the behaviours weaken people's rationality, maintaining the assumption that social acceptance and fear of exclusion for not complying them have more weight.

On the other hand, the second category *(competition between the direct consequences of doing something and the social consequences of rule-following)* includes factors such as:

a) Competition between individual and collective benefit: in this situation, rule following is conditional on the duality between rule following because an immediate individual benefit is received, or because the contingencies will result in long-term collective benefits. In general, it is the common good that prevails favouring the community in general, and so it is necessary to maintain and drive this spirit of cooperation in societies.

b) Motivational tendencies of the subject according to his or her history: when there is a propensity towards social or direct contingencies, as there is an absence of regulation by an authority figure, the possibility of rules being followed reduces. Nonetheless, it must be taken into account that coercive figures or conscious learning processes are not required; it is enough to show what others do in conversations, in advertising, in what is broadcast in the media: something called a *social norm of fulfilment*.

c) Social comparison: people are generally insensitive to rule-following based on perceiving the difficulty of achieving collective behavioural regulation. They deduce that acting individually is ineffective, so they delegate the need to follow rules to other individuals. In consequence, contingencies must be created that promote not only collective actions but also recognitions that extol the importance of individual contributions.

d) Discrepancy in the following of rules between the people who design them or must ensure they are obeyed and those who must follow them (impunity: "don't you know who I am"): this perceived inequality in the consequences of following rules in relation to the consequences for others when following the rule hinders following them. It must be taken into account that what ensures rules are followed is not just the possibility of

being sanctioned but the knowledge that this is considered to be a way of acting shared by other people. If we look and see other people following rules, we will follow them, but if what is observed or divulged is that other people do not follow the rules or that the people who should ensure they are followed break them, then they will not be complied with.

e) The effort involved in rule following: this leads to cumbersome procedures or hinders appropriate following by the people who should ensure they are followed, making it more likely that they will not be followed. This makes necessary to work on better design of rules [a little push (Thaler & Sustein, 2008)] to facilitate following them.

f) Social reproach: although this factor increases the likelihood of people abiding by rules, sometimes, in contrast to what is expected, people are not judged for not following the rule but are condemned socially for following them. Social sanctions are more effective than legal sanctions or individual incentives, and so people must be educated in a culture of recognising following rules as being worth the effort (through incentives) and that disobeying them leads to real sanctions, both legal and social (fines / rejection).

g) Moral compensation: justifying bad actions to avoid taking on behaviours that require greater effort.

This can be illustrated, for example, when people decide to recycle bottles instead of reducing or eradicating consumption of bottled water, thus morally compensating for their behaviour through this small action. In this case it becomes necessary to educate society to be more morally and socially disciplined to generate true regulation.

h) Conflicts between following rules that involve simultaneous behaviours, before the simultaneous action of rules to follow: here following the rule is subject to the greatest aversive consequence or the most positive one. In relation to this, teaching people to evaluate the consequences between rules that come into conflict based on the size of the benefit or sanction received is proposed. Similarly, the importance of establishing some kind of dilemma (moral, social, or legal) for the individual when designing rules is highlighted as this reduces the chance of it being disobeyed.

6. The role of rules in the design of macro- and meta-contingencies for environmental education, mitigating climate change and adapting to it

When educating large population groups about shared problems as in the case of climate change, contingencies must be designed in such a way that they are directed at groups of individuals for which pro-environmental rules must

serve as a bridge in the design of macro-contingencies (MC) and meta-contingencies (MT).

According to Glenn (1988), the former are defined as the agreement between social practices and the results that maintain them; that is to say, the accumulated action of multiple behaviours that produces benefits/harms for all individuals. In this way, "a macro-contingency is established when a functional relationship is created between a social practice that a group of people shares and the result of this collective action as the summation of individual behaviours" (Páramo et al., in press). One example of this could be when people reduce unnecessary use of paper, which, as well as generating social approval for each of the individuals involved in this action, is reflected in the conservation of forests, thus ensuring the cleanliness of the air and the regulation of the climate, benefiting all of society.

The second type, meta-contingencies, are defined as the compromise between individual *interlinked* practices and the accumulated results that maintain them, producing greater benefit for the group than is obtained by individual actions (Glenn, 1991, 2004; Glenn et al., 2016). For example, when a work team whose members know one another directs its efforts to conserving a public park (some taking care of cleaning it, others taking charge of its surveillance, and others still of the conservation of the animals that live in it and the monuments that adorn it), each of these individual contributions will produce a cumulative

effect that benefits both the individual and the group in general. The relationship between the many interdependent behaviours and their consequences leads to a meta-contingency that results in the selection of a group of behaviours or cultural practices and their consequences that the environment provides, guiding people's behaviour and guaranteeing its upkeep by means of cultural transmission of these practices. Weather warnings might be ineffective in the context of metacontingencies that favour practices that are at odds with climate action (Pietras, 2022). In other words, individuals will cooperate with one another when their interdependent behaviour produces greater favourable consequences than independent behaviour (Glenn, 1991, 2003; Glenn et al., 2016).

Accordingly, the rule is the verbal mechanism that establishes relationships of contingency in macro-contingencies and meta-contingencies.

7. Conclusions

This article identifies the importance of directing environmental education towards teaching pro-environmental rules. This approach is based on an epistemological and theoretical position centred on the acquisition and maintenance of language owing to its consequences in a functional analysis that differs from innatism but recognises the continuum between biology and culture.

In recent years, the accumulated evidence has led language theorists to

accept the interaction between the evolutionary process that resulted in the biological structure of the phonological apparatus and the possibility of exercising operant control of it and the functional relationship of language with the social context. As a result, it is accepted that language emerges from the interaction of three different adaptive systems: biological evolution, individual learning, and cultural transmission. This suggests that biological adaptation and cultural transmission have both interacted in the evolution and acquisition of language. Consequently, there is a need to redefine the category of learning, now incorporating the biological antecedents that influence what is learned and the speed with which this process takes place. We conclude that Chomsky's and Skinner's positions are both partially true and complement one another, and so it is clear that the theory of learning must be readjusted to include the biological and environmental aspects of language acquisition, within a concept such as open instinct.

Moreover, it is not necessary to regard learning as a cognitive process, but rather as a complex interaction between our biology, our surroundings, and the relationship that is established between language and people's social verbal setting. One of the conceptual contributions in the article that should be highlighted is that verbal regulation is the cornerstone of what is conceived of as self-regulation, and distances itself from the established outlooks from the cognitive position.

In this way, education in pro-environmental rules becomes the pedagogical device that provides more effective environmental benefits by establishing precisely what is expected for the scope of the desired goals, anticipating by means of various communication and education strategies the insensitivities that usually emerge for following them. It also does this by establishing the macro- and meta-contingencies that tend towards both individual and collective benefit when safeguarding the planet.

Finally, it is important to underline that, other than those that are legal in nature, verbal rules are a guide for action. So, education in pro-environmental rules should not be seen as a mandatory condition or condition of control for the person; they provide guidance of the benefits that would be obtained at both an individual and a collective level when protecting the environment.

Notes

[1] The author participated in this article as part of her studies in the Inter-Institutional Doctoral Program in Education at the Universidad Pedagógica Nacional, Colombia.

[2] Unlike in ordinary language, where the term *contingency* refers to 'uncertainty regarding something', this article uses the definition from behavioral psychology. In this context, contingency is understood as a relationship of dependency between a consequence (whether positive or negative) and a behavior. For example, the benefit will be received whenever the expected behavior occurs.

[3] For enthusiasts, see the critique that Ribes (2000) makes of the inadequate definition of *rule* as a 'discriminative stimulus'.

Authors' contributions

Leidy-Vanessa Díaz-Beltrán: Investigation; Writing (Original Draft Preparation); Visualisation.

Pablo Páramo: Conceptualisation; Supervision; Writing (review and editing); Visualisation.

References

Alcock, J. (2003). *The triumph of sociobiology*. Oxford.

Álvarez-González, C. J. (2010). La relación entre lenguaje y pensamiento de Vigotsky en el desarrollo de la psicolingüística moderna [Thinking and language in Vigotsky and the presence of his theory in the modern psycholinguistics]. *RLA, Revista de Lingüística Teórica y Aplicada*, *48* (2), 13-32. https://dx.doi.org/10.4067/S0718-48832010000200002

Bandura, A., & Walters, R. H. (1959). *Adolescent aggression: A study of the influence of child-training practices and family interrelationships*. Ronald.

Baum, W. M. (2017). *Understanding behaviorism: Behavior, culture, and evolution*. John Wiley & Sons.

Bermúdez, J. (2021). *¿Por qué incumplimos la ley? Carta a un joven estudiante [Why do we break the law? Letter to a young student]*. Ariel.

Broadbent, D. E. (1958). The selective nature of learning [El carácter selectivo del aprendizaje]. In D. E. Broadbent, *Perception and communication [Percepción y comunicación]* (pp. 244-267). Pergamon Press.

Catania, A. C. (2003). Verbal governance, verbal shaping, and attention to verbal stimuli. In K. A. Lattal, y P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 301-321). Springer US.

Cerutti, D. T. (1989). Discrimination theory of rule-governed behavior. *Journal of the experimental analysis of behavior*, *51* (2), 259-276.

Chawla, L., & Derr, V. (2012). The development of conservation behaviors in childhood and youth. In Susan D. Clayton (Ed.), *The Oxford handbook of environmental and conservation psychology* (pp. 527-555). Oxford University Press.

Chomsky, N. (1959). A review of B. F. Skinner's *Verbal behavior*. *Language*, *35* (1), 26-58.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58* (6), 1015-1026.

Corraliza, J. A., & Collado, S. (2011). La naturaleza cercana como moderadora del estrés infantil [Close nature as a moderator of childhood stress]. *Psicothema*, *23* (2), 221-226.

De la Fuente, J. (2017). Theory of self- vs. externally-regulated learning™: Fundamentals, evidence, and applicability. *Frontiers in Psychology, 8*, 1675.

De la Fuente, J., González-Torres, M. C., Aznárez-Sanado, M., Martínez-Vicente, J. M., Peralta-Sánchez, F. J., & Vera, M. M. (2019). Implications of unconnected micro, molecular, and molar level research in psychology: The case of executive functions, self-regulation, and external regulation. *Frontiers in Psychology*, *10*, 1919.

Gifford, R. (2011). The dragons of inaction: psychological barriers that limit climate change mitigation and adaptation. *American Psychologist*, *66* (4), 290-302.

Glenn, S. S. (1988). Contingencies and metacontingencies: Toward a synthesis of behavior analysis and cultural materialism. *The Behavior Analyst*, *11* (2), 161-179.

Glenn, S. S. (1991). Contingencies and metacontingencies: Relations among behavioral, cultural, and biological evolution. In P. A. Lamal (Ed.), *Behavioral analysis of societies and cultural practices* (pp. 39–73). Hemisphere Publishing Corp.

Glenn, S. S. (2003). Selección en dos niveles en la evolución de la ciencia [Two-level selection in the evolution of science]. *Revista Latinoamericana de Psicología*, *35* (3), 281-288.

Glenn, S. S. (2004). Individual behavior, culture, and social change. *The Behavior Analyst*, *27* (2), 133-151.

Glenn, S. S., Malott, M. E., Andery, M. A. P. A., Benvenuti, M., Houmanfar, R. A., Sandaker, I., Todorov, J. C., Tourinho, E. Z., & Vasconcelos, L. A. (2016). Toward consistent terminology in a behaviorist approach to cultural analysis. *Behavior and Social Issues, 25* (1), 11-27.

Gómez, I., Moreno, E., & López, N. (2006). *Sensibilidad a unas u otras contingencias en el marco de la conducta gobernada por regla*s *[Sensitivity to one or the other contingencies in the framework of rule-governed conduct].* Plaza y Valdés.

GramaPsico. (2023, 3 de mayo) *LANGUAGE and VERBAL BEHAVIOR II ~ NOAM CHOMSKY and CHARLES CATANIA #110* [Video]. YouTube. https://www.youtube.com/watch?v=9bwLMXob-AI&t=2299s&ab_channel=enGramaPsico%CE%A8

Gutiérrez, G. (2005). I. P. Pavlov: 100 años de investigación del aprendizaje asociativo [I. P. Pavlov: 100 years of associative learning research]. *Universitas Psychologica*, *4* (2), 251-255.

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Kissi, A. (2020). The study of rule-governed behavior and derived stimulus relations: Bridging the gap. *Perspectives on Behavior Science*, *43* (2), 361-385.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Springer Science & Business Media.

Kissi, A., Hughes, S., Mertens, G., Barnes-Holmes, D., De Houwer, J., & Crombez, G. (2017). A systematic review of pliance, tracking, and augmenting. *Behavior modification*, *41* (5), 683-707.

Lieberman, P. (1998). *Eve spoke: Human language and human evolution*. WW Norton.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior*. Liveright.

MacCorquodale, K. (1969). B. F. Skinner's *Verbal behavior*: A retrospective appreciation. *Journal of the experimental analysis of behavior*, *12* (5), 831-841.

Marler, P., & Terrace, H. S. (Eds.) (1984). *The biology of learning: Report of the Dahlem Workshop on the Biology of Learning, Berlin 1983, October 23-28*. Springer-Verlag.

Midgley, M. (1978). *Beast and man: The roots of human nature*. Cornell University Press.

Neisser, U. (1967). *Cognitive psychology*. Appleton-Century-Crofts.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Páramo, P. (2010). Aprendizaje situado: creación y modificación de prácticas sociales en el espacio público urbano [Situated learning: Creation and modification of social practices in urban public space]. *Psicologia & Sociedade*, *22* (1), 130-138.

Páramo, P. (2013). Comportamiento urbano responsable: las reglas de convivencia en el espacio público [Responsible urban behavior: Rules of conviviality in public space]. *Revista Latinoamericana de Psicología*, *45* (3), 473-485. http://dx.doi.org/10.14349/rlp.v45i3.1488

Páramo, P. (2017). Reglas proambientales: una alternativa para disminuir la brecha entre el decir-hacer en la educación ambiental [Pro-environmental rules: An alternative for reducing the "say-do" gap in environmental education]. *Suma psicológica*, *24* (1), 42-58. https://doi.org/10.1016/j.sumpsi.2016.11.001

Páramo, P., & Burbano, A. (2019). Gramática espacial urbana: la orientación espacial guiada por reglas verbales [Spatial and urban grammar: Spatial orientation focuses on verbal rules]. *Psicogente*, *22* (41), 308-332. https://doi.org/10.17081/psico.22.41.3313

Páramo, P., Burbano, A., Ojeda, G., & Angulo, E. (In press). La señalética urbana orientada a la convivencia [Coexistence-oriented urban signage]. *Revista de Arquitectura (Bogotá)*.

Peláez, M., & Moreno, R. (1999). Four dimensions of rules and their correspondence to rule-governed behavior: A taxonomy. *Behavioral Development Bulletin*, *8* (1), 21-27.

Piaget, J. (1978). *La equilibración de las estructuras cognitivas, problema central del desarrollo [Balancing cognitive structures, a central developmental issue]*. Siglo XXI.

Pietras, C. J. (2022). Rule-governed behavior and climate change: Why climate warnings fail to motivate sufficient action. *Behavior and Social Issues*, *31* (1), 373-417. https://doi.org/10.1007/s42822-022-00109-y

Pinker, S (2007). *The language instinct*: *How the mind creates language*. Harper Perennial Modern Classics.

Pinker, S (2021). *Racionalidad: qué es, por qué escasea y cómo promoverla [Rationality: What it is, why it is scarce and how to promote it]*. Paidós.

Ribes, E. (2000). Instructions, rules, and abstraction: A misconstrued relation. *Behavior and philosophy*, *28* (1-2), 41-55.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107* (6), 2408-2412. https://doi.org/10.1073/pnas.0908239106

Seligman, M. E., & Hager, J. L. (1972). *Biological boundaries of learning*. Appleton Century Crofts.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. B. F. Skinner Foundation.

Skinner, B. F. (1953). *Science and human behavior.* The Free Press.

Skinner, B. F. (1986). The evolution of verbal behavior. *Journal of the Experimental Analysis of Behavior, 45*, 115-122. https://doi.org/10.1901/jeab.1986.45-115

Skinner, B. F. (Ed.) (1981). *Conducta verbal [Verbal behaviour]*. Editorial Trillas.

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin Books.

Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. MIT Press.

Wilson, E. O. (2012). *La conquista social de la Tierra [The social conquest of the Earth]*. Debate.

Wilson, E.O. (2000). *Sociobiology: The new synthesis*. Harvard University Press

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American educational research journal*, *45* (1), 166-183.

Authors' biographies

Leidy-Vanessa Díaz-Beltrán. Inter-institutional doctoral student in Education at the Universidad Pedagógica Nacional. Licentiate degree in Humanities and Languages from the Universidad Libre, master's in Education specialising in English Teaching from the Universidad Externado, Colombia. Her interest in environmental education has led to her obtaining collective awards such as the BIBO 2014 prize with the Escuela Pedagógica Experimental. She currently works as a teacher in the Secretariat of Education, Bogotá (Colombia).

iD https://orcid.org/0009-0004-7731-2077

Pablo Páramo. He earned his PhD in Psychology from City University of New York (CUNY) as a Fulbright scholar and his Master of Science degree from the University of Surrey (England) as a British Council scholar. He obtained his undergraduate degree in Psychology from the Universidad Católica de Colombia. He was recently honoured with the Augusto Angel Maya Award for his contributions to Colombian environmental

thought and has been recognised as an Emeritus Researcher by the Ministry of Science of Colombia. He is currently a faculty member in the Doctorate in Ed- ucation programme at the Universidad Pedagógica Nacional de Colombia.

https://orcid.org/0000-0002-4551-3040

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

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

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

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

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

Abstract:

The question of language in school has an educational dimension of considerable social interest in Spain since a strong process of immersion in regional languages might be affecting students' performance depending on their linguistic and socio-cultural background. In this context, the paper's aim is to analyse the relations between economic, social, and cultural status (ESCS); the percentage of students who speak a different language at school than at home (language) and school performance in PISA (performance); and in particular to measure the predictive power of the second variable with respect to the third, from a territorial perspective. A methodology was developed based on bivariate linear regression analysis and structural equation modelling, and was applied to data regarding these three variables from the PISA 2015 and PISA 2022 databases. These secondary analyses have confirmed: (a) the importance of the predictive power of ESCS on performance consistently in both 2015 and 2022; (b) the intensification of the association between the two variables in that period; and (c) the emergence of a moderate but significant association of the language variable as a predictor of Performance at the territorial level. The impact of regional ESCS differences on performance increased significantly between 2015 and 2022, which is compatible with an increase in other territorial divides described in the literature. Moreover, the strong linguistic immersion models display

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

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

Resumen:

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

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

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

1. Introduction

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1. To provide a comparative overview, by means of a descriptive analysis, of the 2015 and 2022 PISA results, with a detailed presentation of the scores for: performance; Economic, Social, and Cultural Status (ESCS); and language in different autonomous regions and cities of Spain.

2. To construct and evaluate a structural equation model (SEM) of the influence of the ESCS index on academic performance in science, mathematics, and reading on the sample of data from PISA 2015.

3. To analyse the impact of the language spoken at home on academic performance in PISA 2022 by means of a bivariate linear regression analysis.

4. To build and evaluate an SEM (structural equation model) of the relationship between ESCS and the percentage of students who speak a different language to the one that is dominant in schools (language) and its influence on academic performance in science, mathematics, and reading in the sample of data from PISA 2022.

2. Method

2.1. Samples

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

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

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

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

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

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

2.2. Variables

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

Performance latent variable (performance), disaggregated by areas:

- Mean performance in mathematics: this includes the gross values for the mathematics subject evaluated in 2015 (Per_Mat_15) and in 2022 (Per_Mat_22).

- Mean performance in language: this includes the gross values for the language subject evaluated in 2015 (Per_Rea_15) and in 2022 (Per_Rea_22).

- Mean performance in science: this includes the gross values for the science subjects evaluated in 2015 (Per_Sci_15) and in 2022 (Per_Sci_22).

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

- ESCS of the values reported in PISA 2015 (ESCS_15).

- ESCS of the values reported in PISA 2022 (ESCS_22).

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

- Percentage of students who speak a different language at home and at school, according to the figures from the 2022 PISA report (Language_22).

2.3. Analytical procedures

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

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

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

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

3. Results

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

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

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

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

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

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

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

- Regarding the performance variables, with the exception of Asturias, Cantabria, and La Rioja, which improve in at least one of the areas, all of the others worsen in all three. Accordingly, when the mean scores are taken into account, all of them apart from the first two worsen. Catalonia with 31 PISA points and Castile-La Mancha with 25 are the ones that display the largest declines.

- In the case of ESCS, all of the autonomous communities improve in relation to the OECD mean in the period between 2015 and 2022, which results in Spain improving by .48 points, positioning itself above the updated mean. The improvements in Andalusia and the Balearic Islands stand out, with increases of .69 and .60 ESCS points respectively. This phenomenon is analysed in the discussion section.

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

3.2. Structural equations model corresponding to PISA 2015

In a first approach, a model was defined that centred exclusively on the 2015 data, establishing the relationship between ESCS and performance in science, mathematics, and reading. Application of the SEM using the maximum likelihood method with non-linear optimisation evaluated a model that incorporates latent variables (ESCS and performance) and observed variables. This model reveals significant relations between ESCS_15 and academic performance, supported by statistical tests and with an adequate fit.

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

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

	Values obtained
Value of p in $\chi2$	.008
$\chi2/df$	4.87
GFI	1
AGFI	0.999
SRMR	0.047
RMSEA	0.009
CFI	0.91
TLI	0.731
Bentler–Bonett Normed Fit Index (NFI)	0.895
Parsimony Goodness of Fit Index (PGFI)	0.143
Parsimony Normed Fit Index (PNFI)	0.298
Akaike Criterion (AIC)	378.019

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

Similarly, Table 4 shows the estimates for the model for measuring the latent variables. ESCS_15, as the observed variable of the ESCS latent variable, has a standardised coefficient of 1.000, which indicates a perfect relationship. Likewise, with regards to academic performance, it can be seen that the observed variables (Per_Sci_15, Per_Mat_15, Per_Rea_15) have standardised coefficients of 1.00, 0.99, and 0.90 respectively, signalling the relationships existing between these observed variables and the latent variable of performance.

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

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

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

Figure 1 presents the path model illustrating the impact of ESCS on performance in PISA 2015. The SEM with data from PISA 2015 underlines the importance of ESCS in predicting academic performance, supporting the hypothesis that higher ESCS is positively linked to better performance in science, mathematics, and reading.

FIGURE 1. Path model of the relationship between ESCS and academic
performance in PISA 2015.
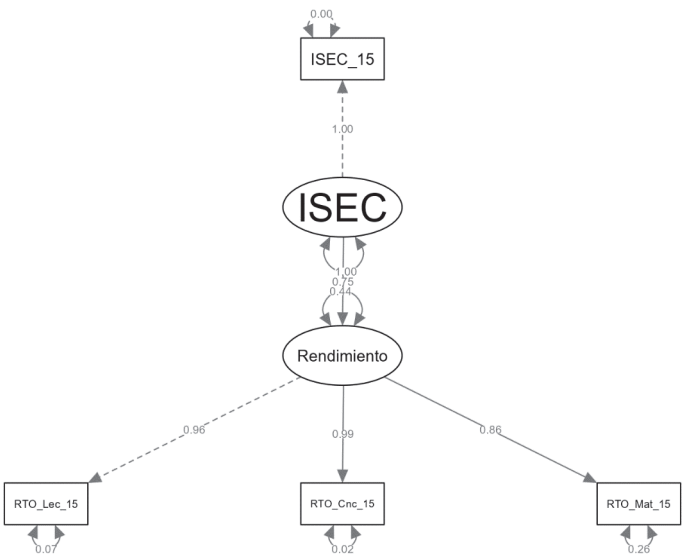
3.3. The predictive value of the language variable with regards to performance in PISA

A bivariate linear regression analysis was performed to analyse the impact of the language variable on academic performance in PISA 2022. Table 5 presents key metrics that illustrate the connection between the two variables. The multiple correlation coefficient displays a weak relationship, while the low R^2 values indicate that only a small proportion of the variability in performance is attributed to language.

TABLE 5. Regression statistics.

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

Table 6 shows the global evaluation of the model in greater depth. The lack of statistical significance ($F = 0.59$, $p > .05$) and the modest sum of squares in the regression point to a generally weak direct relationship between language and performance. Furthermore, the coefficients associated with the intercept and the language variable did not achieve sufficient significance.

TABLE 6. Analysis of variance and coefficients.

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

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

Although the Language variable does not show a statistically significant impact on the PISA 2022 performance scores, graphical analysis of this weak relationship (see Figure 2) shows an inverse trend between the two variables: performance in PISA 2022 and use of language in samples at OECD, national, and sub-national levels.

FIGURE 2. Linear regression analysis of PISA 2022 performance vs language on the OECD, national, and sub-national samples.
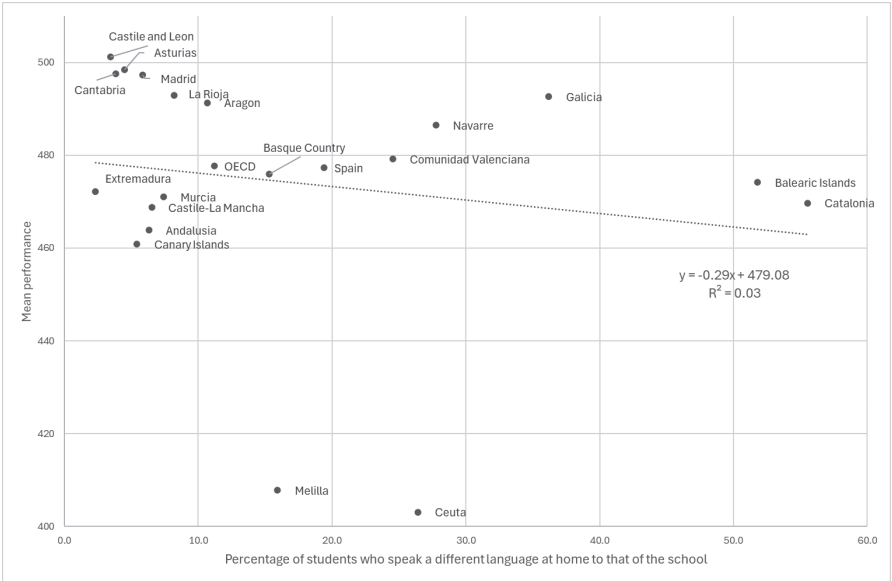
3.4. Structural equation model corresponding to PISA 2022

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

The application of the SEM analysis using the maximum likelihood method with non-linear optimisation has evaluated a model that includes latent variables (ESCS, language, and performance) and observed variables. This model finds significant relationships between ESCS_22 and academic performance, as well as between the percentage of students who speak a different language to that of the schools (language), supported by statistical tests and an adequate fit.

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

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

	Values obtained
Value of p in $\chi2$	.01
$\chi2/df$	2.65
GFI	1
AGFI	0.999
SRMR	0.04
RMSEA	0.014
CFI	0.939
TLI	0.869
NFI	0.909
PGFI	0.259
PNFI	0.424
AIC	642.657

Table 8 shows the parameter estimates as well as the measurement model for the relationship between academic performance, language, and ESCS in 2022. In the section relating to academic performance, an increase of 1 ESCS unit is seen to be associated with an estimated increase of 166.78 units in academic performance, with a confidence interval of 95% between 98.78 and 234.78. This coefficient is significant ($p <.001$), indicating a positive relationship between ESCS and academic performance. Likewise, for the language predictor, it is found that an increase of 1 unit in the percentage of students who speak a language other than that of the school is related with an estimated reduction of 0.43 units in academic performance, with a confidence interval of 95% between -0.81 and -0.04. This coefficient is also significant ($p <.05$), signalling a negative relationship between language and performance.

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

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

Latent variable	Observed variable							
ESCS	ESCS_non_immig_22	1.00	0.00	1.00	1.00	0.805		
	ESCS_im-mig_22	1.53	0.38	0.78	2.28	0.723	3.99	<.001
Language	%_languag-es_22	1.00	0.00	1.00	1.00	1.000		
	Per_Mat_22	1.00	0.00	1.00	1.00	0.977		
Performance	Per_Rea_22	0.92	0.05	0.82	1.02	0.992	18.18	<.001
	Per_Sci_22	0.94	0.05	0.85	1.05	0.995	19.09	<.001

With regards to the latent and observed variables, it should be noted that the relationships between ESCS and its components, as well as between language and the percentage of students who speak different languages, are significant and coherent with the model proposed. The observed variables of performance in the specific subjects also display positive and significant relationships with the latent performance variable. As a whole, these detailed estimates offer a deeper understanding of the interactions between the key variables in the context of the SEM model applied to the year 2022.

In summary, regarding the specific relationships from PISA 2022, a positive and strong association has been observed between ESCS and performance, as is a negative relationship of moderate but significant intensity between language and academic performance. These conclusions are supported visually by the path model in Figure 3.

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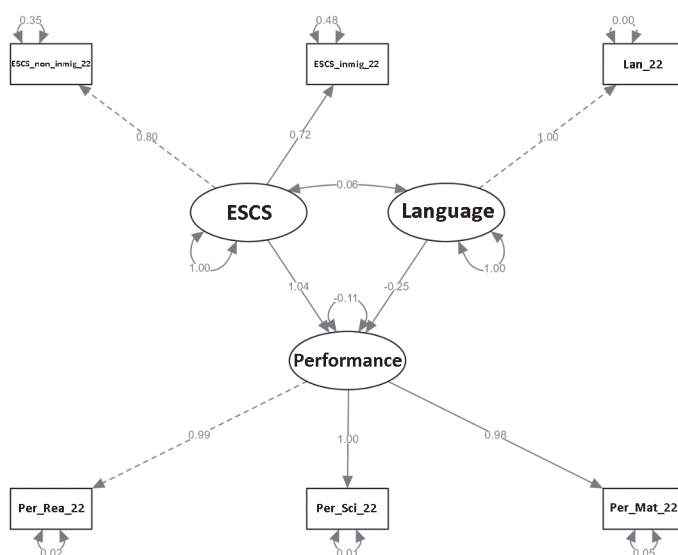
4. Discussion

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

4.1. The influence of ESCS on performance

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

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

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

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

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

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

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

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

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

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

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

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

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

In addition to the statistically significant relations between ESCS_22 and academic performance, the model reveals the connection between this and the percentage of students who speak a language other than that of the school at home (language). These associations support the idea that socio-economic context and linguistic variables both contribute significantly to student performance.

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

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

5. Conclusions

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

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

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

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

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

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

Authors' contributions

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

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

Isabel García-García: Data curation; Visualisation.

References

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, *193*, 104345. https://doi.org/10.1016/j.jpubeco.2020.104345

Ball, J. (2011). *Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years*. United Nations Educational, Scientific and Cultural Organization (UNESCO). https://unesdoc.unesco.org/ark:/48223/pf0000212270

Benson, C. (2004). *The importance of mother tongue-based schooling for educational quality*. United Nations Educational, Scientific and Cultural Organization (UNESCO). https://unesdoc.unesco.org/ark:/48223/pf0000146632

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16* (4), 240-250.

Bryman, A., & Cramer, D. (1994). *Quantitative data analysis for social scientists*. Routledge.

Bühmann, D., & Trudell, B. (2007). *Mother tongue matters: Local language as a key to effective learning*. United Nations Educational, Scientific and Cultural Organization (UNESCO). https://unesdoc.unesco.org/ark:/48223/pf0000161121

Bulut, O., Delen, E., & Kaya, F. (2012). An SEM model based on PISA 2009 in Turkey: How does the use of technology and self-regulation activities predict reading scores? *Procedia - Social and Behavioral Sciences*, *64*, 564-573. https://doi.org/10.1016/j.sbspro.2012.11.066

Calero, J., & Choi, Á. (2019). *Efectos de la inmersión lingüística sobre el alumnado castellanoparlante en Cataluña [Effects of linguistic immersion on Spanish-speaking pupils in Catalonia]*. Fundación Europea Sociedad y Educación. https://sociedadyeducacion.org/site/wp-content/uploads/SE-Inmersion-Cataluna.pdf

Carabaña, J., & Fernández-Enguita, M. (2019). Enseñanza en lenguas regionales y sentimiento de pertenencia a la escuela [Schooling in regional languages and sense of belonging at school]. *Journal of Supranational Policies of Education*, (9), 57-94. http://doi.org/10.15366/jospoe2019.9.002

Castro, M., & Egido, I. (2024). *Qué sabemos sobre el profesorado. Políticas, evidencias y perspectivas de futuro [What we know about teachers. Policies, evidence and future perspectives]*. Narcea.

Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). How did covid-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data. *NBER Working Paper*, (27431). https://doi.org/10.3386/w27431

Clots-Figueras, I., & Masella, P. (2013). Education, language and identity. *The Economic Journal*, *123* (570), 332-357.

Cummins, J. (2001). Bilingual children's mother tongue: Why is it important for education? *Sprogforum,7* (19),15-20.

Epskamp, S., Stuber, S., Nak, J., Veenman M., & Jorgensen T. D. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output* [Computer Software]. https://CRAN.R-project.org/package=semPlot

Foguet, J. M. B., & Gallart, G. C. (2000). *Modelos de ecuaciones estructurales [Structural equation modelling]*. La Muralla.

Gairín, J. (2024). *Dirección y liderazgo de los centros educativos. Naturaleza, desarrollo y práctica profesional [School management and leadership. Nature, development and professional practice]*. Narcea.

Gallucci, M., & Jentschke, S. (2021). *SEMLj: jamovi SEM Analysis* [Computer Software]. https://semlj.github.io/

Hattie, J. (2023). *Visible learning: The sequel. A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge.

Hovens, M. (2002). Bilingual education in West Africa: Does it work? *International Journal of Bilingual Education and Bilingualism, 5* (5), 249-266.

Kim, J. O., & Kohout, F. J. (1988). Special topics in general linear models. In N. H. Nie, D. H. Bent, & C. H. Hull, *Statistical package for the social sciences*. McGraw- Hill.

López-Rupérez, F., García-García, I., & Expósito Casas, E. (2018). *PISA 2015 y las comunidades autónomas españolas. Diagnósticos y políticas de mejora [PISA 2015 and the Spanish autonomous communities. Diagnostics and improvement policies]*. Universidad Camilo José Cela. https://www.ucjc.edu/pdfs/universidad/pisa2015_ccaa.pdf

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019a). Rendimiento en ciencias, concepciones epistémicas y vocaciones STEM en las comunidades autónomas españolas. Evidencias desde PISA 2015, políticas y prácticas de mejora [Performance in science, epistemic conceptions, and STEM vocations in Spain's autonomous communities: evidence from PISA 2015, improvement policies, and practices]. **Revista Española de Pedagogía**, *272* (77), 5-27. https://doi.org/10.22550/REP77-1-2019-09

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019b). Educational effectiveness, efficiency, and equity in Spanish regions: What does PISA 2015 reveal? *ORBIS SCHOLAE*, *12* (2), 9-36. https://doi.org/10.14712/23363177.2018.291

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2019c). Igualdad de oportunidades e inclusión educativa en España [Equal opportunities and educational inclusion in Spain]. *RELIEVE*, *25* (2), 1. http://doi.org/10.7203/relieve.25.2.14351

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España: análisis empíricos y recomendaciones políticas [Grade repetition and graduation in compulsory secondary education in Spain: Empirical analysis and policy recommendations]. *Revista de Educación*, *394*, 325-354. https://doi.org/10.4438/1988-592X-RE-2021-394-510

López-Rupérez, F., & García-García, I. (2023, 11 Decembre). PISA 2022 y la cuestión territorial en España [PISA 2022 and the territorial question in Spain]. *Magisnet*. https://www.magisnet.com/2023/12/pisa-2022-y-la-cuestion-territorial-en-espana/

OECD. (2016). *PISA 2015 results (volume I). Excellence and equity in education*. OECD Publishing. https://doi.org/10.1787/9789264266490-en

OECD. (2023). *PISA 2022 results (volume I). The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Parliamentary Assembly. (2006). *The place of mother tongue in school education* (Doc. 10837). Committee on Culture, Science and Education. Council of Europe. https://assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewHTML.asp?FileID=11142&lang=EN

Pearl, J., y Mackenzie, D. (2020). El libro del porqué. *La nueva ciencia de la causa y el efecto [The book of why. The new science of cause and effect]*. Pasado y Presente.

Pérez, E., Medrano, L. A., & Sánchez-Rosas, J. (2013). El *path analysis*: conceptos básicos y ejemplos de aplicación [Path analysis: Basic concepts and examples]. *Revista Argentina de Ciencias del Comportamiento, 5* (1), 52-66.

Pinnock, H. (2009a). *Language and education. The missing link. How the language used in schools threatens the achievement of education for all*. CfBT Education Trust & Save the Children. https://resourcecentre.savethechildren.net/pdf/1674.pdf/

Pinnock, H. (2009b). *Steps towards learning: A guide to overcoming language barriers in children´s education*. Save the Children. https://www.savethechildren.org.uk/content/dam/global/reports/steps-towards-learning-lr.pdf

Reilly, C. (2019). *Language and development: Issues in mother tongue based multilingual education* [Inedit Manuscript]. https://www.researchgate.net/publication/343768795_Language_and_development_Issues_in_mother-tongue_based_multilingual_education#fullTextFileContent

Rodríguez, P. (2023, 7 Diciembre). Catalunya se pregunta por qué se ha hundido más que el resto de España en PISA [Catalonia wonders why it has sunk further than the rest of Spain in PISA]. *Eldiario.es*. https://www.eldiario.es/catalunya/catalunya-pregunta-hundido-resto-espana-pisa_1_10748162.html

Rodríguez-Sanmartín, O. (2023, 19 Diciembre). Los alumnos de Cataluña son los que muestran el mayor rechazo a la escuela de toda España [Pupils in Catalonia show the highest level of rejection of school in Spain]. *El Mundo*. https://www.elmundo.es/espana/2023/12/19/6581fb-22fdddff570d8b45b5.html

Rosseel, &. (2012). Iavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48* (2), 1-36. https://doi.org/10.18637/jss.v048.i02

Salas, A. A., Navarro, R., & Montero, E. (2017). Un modelo de ecuaciones estructurales para el estudio de factores que afectan la competencia lectora y la alfabetización matemática: una aproximación bayesiana con datos de PISA 2009 [A structural equation model for the study of factors affecting reading literacy and mathematical literacy: A Bayesian approach using PISA 2009 data]. *Estadística Española*, *59* (194), 167-192.

Sanz, I., & Tena, J.D. (2023). The impact of the covid-19 pandemic on education learning. In J. Sanz, & I. Sanz (Eds.), *Addressing inequities in modern educational assessment* (pp. 15-36). Springer. https://doi.org/10.1007/978-3-031-45802-6_2

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

Soler, A. (2022). Heterogeneidad regional en el abandono educativo temprano [Regional heterogeneity in early school dropout rates]. *Magisterio*, 24-25.

Swe, K.M., Fraser, B. J., & Afari, E. (2020). Structural relationships between learning environments and students' non-cognitive outcomes: Secondary analysis of PISA data]. *Learning Environments Research*, *23* (3), 395-412. https://doi.org/10.1007/s10984-020-09313-2

The jamovi project. (2022). *jamovi* (version 2.3) [Computer Software]. https://www.jamovi.org

Thomas, W. P., & Collier, V. P. (1997). *School effectiveness for language minority students*. National Clearinghouse for Bilingual Education. https://ncela.ed.gov/sites/default/files/legacy/files/rcd/BE020890/School_effectiveness_for_langu.pdf

Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Center for Research on Education, Diversity and Excellence. https://escholarship.org/uc/item/65j213pt

UNESCO. (1953). *The use of the vernacular languages in education*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO).

UNESCO. (2016). *Si no entiendes, ¿cómo puedes aprender? Informe de seguimiento de la educación en el mundo* [If you don't understand, how can you learn? Global education monitoring report]. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). https://unesdoc.unesco.org/ark:/48223/pf0000243713_spa

UNESCO-IIEP. (2021, 23 de septiembre). *La lengua de enseñanza [The language of teaching]*. https://bit.ly/3UsQE2C

Woessmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., & Zierow, L. (2020). Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen? [Education during the coronavirus crisis: How have schoolchildren spent the period of school closures and what educational measures are Germans in favour of?]. *ifo Schnelldienst*, *9*, 25-39.

rep

Authors' biographies

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

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

Álvaro Moraleda-Ruano. He has a Doctorate in Pedagogy from the Universidad Complutense de Madrid, with special doctoral prize. His contribution to the Department of Theory and History of Education is reflected in various collaboration grants, including an honorary grant. He has focussed his work in key areas such as emotional intelligence and statistical research methods in education. He is currently research coordinator in the Faculty of Education of the Universidad Camilo José Cela. He has an extensive record of scientific publications and has participated in various externally funded research groups and projects.

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

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

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

Contribution of machine learning to the analysis of grade repetition in Spain: A study based on PISA data

Contribución del machine learning *al análisis de la repetición escolar en España: un estudio con datos PISA*

Alexander CONSTANTE-AMORES. Associate Lecturer. Universidad Camilo José Cela, Spain (*israelalexander.constante@ucjc.es*).

Delia ARROYO-RESINO, PhD. Lecturer. Universidad Complutense de Madrid, Spain (*garroy01@ucm.es*).

María SÁNCHEZ-MUNILLA, PhD. Research Assistant. Universidad Complutense de Madrid, Spain (*marsan27@ucm.es*).

Inmaculada ASENSIO-MUÑOZ, PhD. Senior Lecturer. Universidad Complutense de Madrid, Spain (*macu@edu.ucm.es*).

Abstract:

Introduction: The rate of grade repetition is excessively high in Spain despite being a controversial measure. In order to obtain evidence to contribute to reducing it in compulsory education, the present work is an in-depth study of the PISA 2018 context indices that are most closely linked to this phenomenon. Method: With the sample of Spanish students ($n = 35\,943$), we used an automatic machine learning method to select and order the predictors, and multilevel logistic regression (students and centres) to quantify the contribution of each one. Results: For each educational stage we obtained the 30 most significant contextual variables, which explain 65.5% of the grade repetition variance in primary education and almost 55.7% in secondary education. Conclusions: The main indicators are principally at student level, which suggests the suitability of psychoeducational interventions based more on individual support than general policies. This gives rise to potentially more efficient and equitable measures than grade repetition, aimed at, for example, the management of learning time or academic/professional guidance, and predictors with specific differential significance at each stage. Methodologically, the study contributes to improving the specification of predictive models.

Keywords: PISA, grade repetition, machine learning, contextual variables, multilevel logistic regression, compulsory education.

Resumen:

Introducción: La repetición de curso tiene una excesiva incidencia en España a pesar de ser una medida controvertida. A fin de obtener evidencias que contribuyan a su reducción en la educación obligatoria, el presente trabajo profundiza en el estudio de los índices de contexto de PISA 2018 más vinculados con dicho fenómeno. Método: Con la muestra de estudiantes españoles ($n = 35\,943$), se utiliza un método de aprendizaje automático para seleccionar y ordenar los predictores, y una regresión logística multinivel (estudiantes y centros) para cuantificar la contribución de cada uno. Resultados: Se obtienen las 30 variables de contexto más relevantes en cada etapa edu-

cativa, que explican el 65.5 % de la varianza de la repetición en primaria y casi el 55.7 % en secundaria. Conclusiones: Los principales indicadores son sobre todo del nivel de estudiantes, lo que sugiere la idoneidad de intervenciones psicoeducativas basadas en el apoyo individualizado más que en políticas generalizadas. De ahí emergen medidas potencialmente más eficientes y equitativas que la repetición, centradas, por ejemplo, en la gestión del tiempo de aprendizaje o en la orientación académico-profesional, así como predictores con importancia específica diferencial en cada etapa. En el ámbito metodológico, el estudio hace una aportación a la mejora de la especificación de los modelos predictivos.

Palabras clave: PISA, repetición escolar, aprendizaje automático, variables de contexto, regresión logística multinivel, educación obligatoria.

1. Introduction

As is generally known, grade repetition is a strategy that consists of retaining the student at a level of a specific educational stage when they have not proved that they have acquired the minimum levels of knowledge (Jimerson & Ferguson, 2007; López & García, 2020). Nevertheless, it is a controversial measure, that some authors consider to be positive because it helps lower-performing students to improve their skills and gain maturity (Battisttin & Schizzerotto, 2019; Ikeda & García, 2014; Valbuena et al., 2021) and that others see as a mere-

ly disciplinary measure (Schwerdt et al., 2017). Furthermore, although this issue has been the subject of synthesis studies, systematic reviews and meta-analysis covering practically the entire twentieth century (Allen et al., 2009; Jimerson 2001), there are doubts as to its effectiveness. Moreover, it even seems to be detrimental, as it is associated with a decrease in student performance (Asensio et al., 2018; Hattie, 2017; López et al., 2018, Nieto-Isidro & Martínez-Abad, 2023) and is linked to negative aspects, such as frustration (López-Rupérez et al., 2021), feelings of discrimination

(Van Canegem et al., 2021) or a lowering of self-confidence, academic self-concept or motivation, thus affecting academic expectations (González-Nuevo et al., 2023; Mathys et al., 2019; Peixoto et al., 2016).

The corresponding regulations in Spain take this situation into account and the Organic Law 1/1990 of 3 October, on the General Organisation of the Educational System (LOGSE) established that repetition is a measure that focuses on exceptional diversity. Although, in fact, its use is not so exceptional. As reported by the Ministerio de Educación y Formación Profesional [Ministry of Education and Vocational Training] (2022), in the first stage of secondary education the repetition rate is 8.7% and in primary education 2.4% of students have repeated a year, a percentage that is only surpassed in the European Union by Portugal (3.6%), the Slovak Republic (2.9%) and Austria (2.9%). Furthermore, based on data from the Programme for International Student Assessment (PISA) for 2009, 2015 and 2018, López and García (2020), who compare the percentage of 15-year-old students who have repeated at least once in the European Union, conclude that Spain, despite having considerably reduced the rate between 2009 and 2018, in the latter year still has a rate of 28.71%. That puts the country in third place, only behind Luxembourg (32.02%) and Belgium (29.06%), when the average for the Organisation for Economic Co-operation and Development (OECD) stands at 11.4% and for Finland it is 3.2%.

In view of the above, further research into this phenomenon is necessary, with the ultimate aim of providing information that makes it possible to act based on the evidence, not only from a political perspective, but also through educational practice, thereby contributing to data optimisation so that Spain can get closer, at least, to the OECD average.

To this effect, the present study takes the line of a search for predictors by means of secondary exploitation of the PISA data, for which there is already research based on previous PISA publications, showing that grade repetition has many causes (Arroyo et al., 2019; Carabaña, 2013; Cordero et al., 2014).

In line with the classification of background questionnaires used by PISA (OECD, 2019), shown in Figure 1 below, there is a brief summary of the state of affairs regarding the most significant variables of the three constructs.

Regarding precedents, there is evidence that, above all, the students who repeat grades usually come from underprivileged socio-economic backgrounds (López & García, 2020). However, immigration background shows inconclusive results, as Cordero (2014), García-Pérez et al. (2014) o Warren et al. (2014) indicate that first-generation immigrants (born abroad) have a greater probability of repeating a grade than natives, while, for example, in the study by Choi et al. (2018) this variable does not prove to be significant.

FIGURE 1. Modular structure of contextual variables in PISA 2018.

	Student background constructs	Schooling constructs			Non-cognitive and metacognitive constructs
		TEACHING AND LEARNING			
		1. Teacher qualifications and knowledge	2. Teaching practices for science	11. Learning time and curriculum	4. Reading related outcomes: attitudes, motivation and strategies
	5. Out-of scool scientific experience		SCHOOL POLICIES		
			3. School-level learning environment for reading		
		13. Implicación de los padres	12. School climate: interpersonal relationships, trust and expectations	14. School context and resources	9. Dispositional and school-focused variables
	8. Educational pathways in early childhood		GOVERNANCE		
			16. Student assessment, institutional assessment and accountability	15. Allocation, selection and choice	10. Dispositions for global competence
	6. Socio-economic status of student and family				
	7. Migration and culture				
Reading literacy					
General categories					

Source: OECD (2019).

Secondly, among the school-related constructs, studies have mainly been conducted on learning time, school climate or the curriculum, which are associated with grade repetition (OECD, 2019). In this way, Asensio et al. (2018) and López et al. (2018) demonstrate the importance of the number of classes per week or the duration of classes for grade repetition, while Seabra and Ferrão (2016) find a significant relationship between grade repetition and disciplinary problems in schools, with the disciplinary climate having a greater impact than the socio-economic level. On the other hand, Arroyo et al. (2019) link repetition with curriculum-related variables, such as having studied science the previous year or having chosen an optional subject in this field.

In the group of non-cognitive and metacognitive constructs (hereinafter NC+MCC), motivation, self-confidence, future academic expectations and a sense of belonging to the group are highly significant variables (Fernández-Lasarte et al., 2019; Hornstra et al., 2017; Van Canegem et al., 2021), so that positive development of these aspects minimises the effect of grade repetition (Marsh et al., 2018). Likewise, academic self-concept and focusing on goals provide a high level of predictive capacity regarding the phenomenon under study (Ferla et al., 2009; Rodríguez-Rodríguez, 2022).

Despite the amount of existing research, one limitation that makes it difficult to define a more accurate picture lies in the fact that predictors are usually introduced into models in a random way,

perhaps due to the lack of theoretical references that, in education, would produce modelling that aims to minimise specification error. Although the main assumption of multilevel models (which are the most widely used in secondary exploitation of PISA data, due to their isomorphism with the reality they aim to model) is that they are well defined, that is to say, that they meet the condition of "not excluding major predictors from the model" (Gaviria & Castro, 2005, p. 86).

In view of this, it is interesting to explore predictive models in which statistical error reduction is matched by the attempt to minimise specification error as well. The present research takes this line, as do others that use data mining, based on machine learning (Urbina & De la Calleja, 2017), both in international studies (Gamazo & Martínez-Abad, 2020; Kılıç et al., 2017; Kiray et al., 2015; Liu & Ruiz, 2008; Martínez-Abad & Chaparro, 2017; Martínez-Abad et al., 2020), and those focusing on Spanish samples (Arroyo et al., 2024a; Arroyo et al., 2024b; Asensio et al., 2018).

Another limitation is that the previous research does not tend to distinguish between students from the two compulsory stages, with most studies concentrating on secondary school students, which prevents the appearance of differentially significant variables in primary education.

In order to contribute to overcoming the aforementioned limitations, the overall objective pursued by this research is to obtain pertinent information for modelling

that is a closer fit for grade repetition in the two compulsory stages in Spain, based on the PISA 2018 databases (Ministerio de Educación y Formación Profesional, 2019), which provide a vast amount of information about contextual variables (OECD, 2019).

The specific goals are:

- To identify the most accurate machine learning algorithm [individual decision trees (CART and C5.0), random forest and stochastic gradient boosting (GBM)] and, with it, select the variables that are most closely linked to grade repetition in the two stages and sort them according to their differential significance in each stage.

- To determine the relative contribution of each factor (student background, school-related and NC+MCC constructs) to the probability of a student repeating a grade in primary and secondary education taking into account the two-level structure of the data (students and schools).

2. Method

A secondary analysis was conducted of the PISA 2018 international assessment data, using a non-experimental, quantitative, cross-disciplinary, correlational design, in accordance with the PISA manual on technical guidelines (OECD, 2019).

2.1. Participants

Using two-stage sampling, the sample of Spanish students who participated in PISA 2018 is 35943, which represents coverage of 92% of the national population of 15-year-olds (Ministerio de Educación y Formación Profesional, 2019), composed of 49.96% women and with an average age of 15.84 years old ($SD = 0.29$). By type of school, 65% are students who are enrolled in public centres, 28% in state-contracted, privately-owned centres and 7% in private schools. The full information can be found in the second annex of the document published by the OECD (2020).

2.2. Instruments

All the variables have been measured with the PISA 2018 background questionnaires for students and principals (OECD, 2019). The dependent variables are grade repetition in primary and secondary education, with almost double the rate in the latter stage as reported by the participants themselves (Table1).

In terms of the independent variables, 85 were introduced in the initial stage (see Tables S1 and S2 at https://acortar.link/hfDnvO). Of this set of predictors, 83 are composite indices (WLE and SUM), of which 67 indices were constructed based on student responses (level 1) and 17 based on the school principals' responses (level 2). The two remaining variables are simple items that have been included in the model due to the large amount of previously existing evidence: *student gender* (level 1) and *percentage of 15-year-old students from socio-economically disadvantaged homes* (level 2).

TABLE 1. Grade repetition in compulsory education in Spain.

Have you ever repeated a grade in primary education?				
	n	%	*n* (weighted)	% (weighted)
No	30 078	89.927	340 287	88.606
Yes, once or more	3369	10.073	43 756	11.394
Total valid responses	33 447			
Have you ever repeated a grade in secondary education?				
	n	%	*n* (weighted)	% (weighted)
No	27 902	80.354	310 497	77.313
Yes, once or more	6822	19.646	91 114	22.687
Total valid responses	34 724			

Source: author's own compilation based on the PISA (2018) data for Spain.

Subsequently, the multilevel models included only those that, using the more accurate machine learning algorithm, show a greater connection with grade repetition in each of the stages studied.

2.3. Procedure

A detailed summary of the procedure for applying the questionnaires can be found in the Spanish report (Ministerio de Educación y Formación Profesional, 2019).

2.4. Data analysis

There were two separate stages, so that first an exploratory, descriptive study was conducted (Rosenthal & Rosnow, 2008) using algorithms supervised by machine learning and, once the most significant variables had been selected and sorted, several predictive models were tested taking into account the hierarchical data structure (level 1: student and level 2: school).

Prior to the machine learning algorithm, the data were pre-processed, as recommended by Raschka (2015), by processing the missing values with the technique of multiple imputation by chained equations (Rivero, 2011) and proceeding to random division of the cases into two sets: one for training (70%), with which to adjust the model, and the other for validation (30%), with which to evaluate the model by means of the area under the curve (AUC), as these are two dependent variables with unbalanced levels. Centring and scaling were also performed on the continuous predictors in order to avoid overrepresentation of the variables with a greater magnitude, and no predictor with zero or near-zero variance was obtained.

To meet the first objective, we compared the performance of four algorithms [individual decision trees (CART and C5.0), random forest and stochastic

rep

gradient boosting] using the AUC, with the grid-search method based on 10-fold cross-validation, jointly optimising the most significant hyperparameters. Using the sorting method of relative importance, Stochastic Gradient Boosting proved to be the most accurate algorithm, in accordance with the Friedman test applied to the training sample. For the sake of parsimony, given that one of the limitations of this algorithm is that it does not suggest a fixed cut-off point regarding the selection of an optimal number of predictors (Sarkar et al., 2018), in line with Gorostiaga and Rojo-Álvarez (2016), we decided to examine the following sets of variables: 15, 20, 25, 30 and 35. In this way, greater precision was provided for the selection of the optimal set, which proved to be, in both primary and secondary education, the one containing 30 of all the indices initially included.

To achieve the second objective, we used the technique of multilevel binary logistic regression, as the random variance of the null model proved to be statistically significant and the intraclass correlation coefficient (ICC) was greater than 10% (Lee, 2000). In considering the assumption of multicollinearity, we eliminated *home possessions* (HOMEPOS), *mother's education* (MISCED) and *parents' highest occupational status* (HISEI), as they are implicit in the *index of economic, social and cultural status* (ESCS); and the predictor *family wealth*, due to its strong correlation with this index.

Following this, as required by multilevel analysis, we removed the schools with fewer than 20 subjects, thereby excluding 1532

subjects belonging to 133 centres, so the final sample consisted of 34 411 students.

This generated a total of five models: the null model, without predictors; model 1, composed of the student background variables; model 2, comprising the school-related variables; model 3, with the NC+MCC variables; and model 4, containing all the predictors. Thus, it was possible to analyse the contribution of the school-related and NC+MCC factors once the student background had been checked. The variables were introduced into the model in the order obtained from the machine learning algorithm (Arroyo et al., 2019; Constante-Amores et al., 2021). Finally, to check the fit of the models, we used the AIC and BIC indices and Deviance, a statistical reduction that allows nested models to be compared by estimating the percentage of reduced variance (R2), and we calculated their significance (Cameron & Windmejier, 1997). The interpretation of the results is based on the percentage of explained variance (PEV), which quantifies the explanatory capacity of the final model, and on the odds ratios, in which values greater than 1 indicate that the variable increases the probability of repeating a grade, acting as a risk factor, and values below 1 suggest that this probability decreases, acting as a protective factor.

In terms of software, we used the version 4.4.0 R packages caret, stochastic gradient boosting and lme4 to compare the machine learning algorithms, select the most important variables and conduct multilevel binary logistic regression, respectively.

3. Results

Of the algorithms tested, the most accurate were stochastic gradient boosting and random forest, for both primary education (Figure 2) and secondary education (Figure 3). With the Friedman test, of note is stochastic gradient boosting, which we eventually used to select and hierarchise the most significant variables associated with grade repetition in compulsory education.

FIGURE 2. Comparison of machine learning algorithms for primary education.
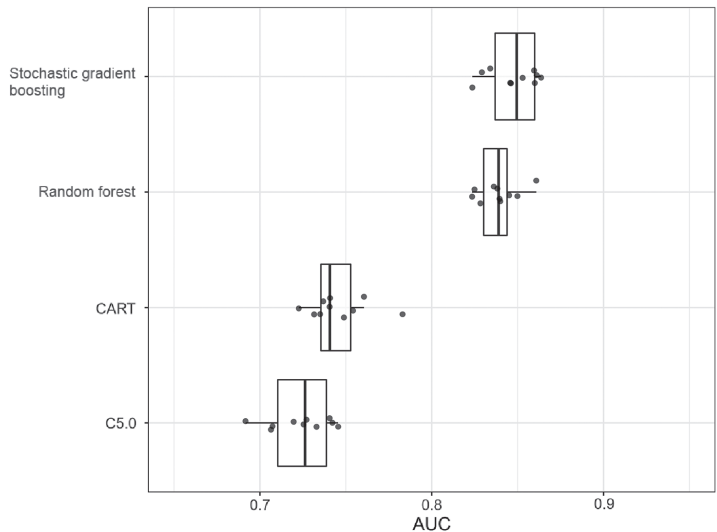

FIGURE 3. Comparison of machine learning algorithms for secondary education.
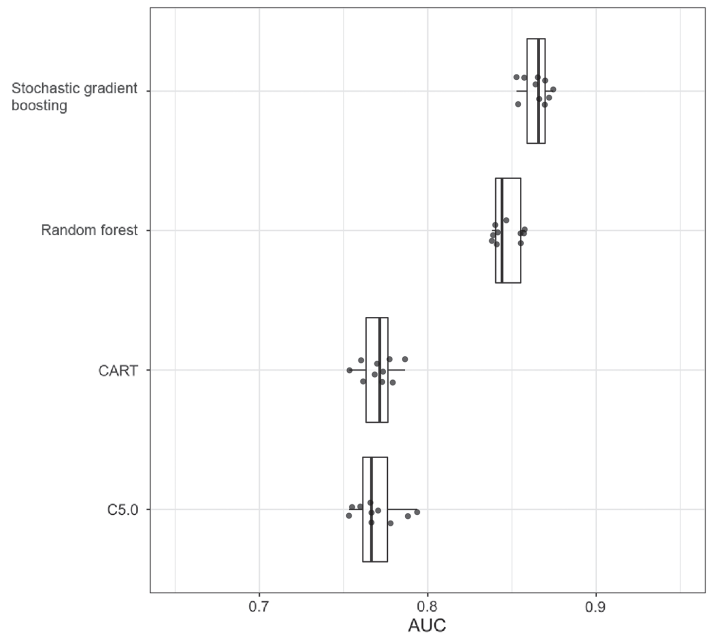
547 EV

Table 2 shows the empirical results produced by stochastic gradient boosting, in which, out of the 30 most significant variables in primary education, 12 relate to background, 9 are associated with the school and 9 are NC+MCC. All of them correspond to level 1, except *the percentage of students from socio-economically disadvantaged homes*, which is from level 2, ranked 23rd. In this educational stage, the most significant predictors are the following: first, the *index of economic, social and cultural status* (ESCS); secondly, the *student's expected occupational status* (BSMJ); and thirdly, *learning time for mathematics* (MMINS).

TABLE 2. Most significant variables associated with grade repetition in primary education.

	Label	Order	Variables
Student background	ESCS	1st	Index of economic, social and cultural status
	HOMEPOS	4th	Home possessions
	SCCHANGE	5th	Number of school changes
	HISEI	8th	Parents' highest occupational status
	MISCED	9th	Mother's education
	IMMIG	12th	Immigration background
	ICTRES	16th	Home ICT resources
	HEDRES	17th	Home educational resources
	DURECEC	20th	Duration of early childhood education and care
	SC048Q03NA	23rd	Percentage of students from socio-economically disadvantaged homes
	BMMJ1	27th	Mother's ISEI
	WEALTH	30th	Family wealth
Related to school	MMINS	3rd	Learning time for mathematics (minutes per week)
	LMINS	6th	Learning time for language (minutes per week)
	INFOCAR	10th	Information about careers
	SMINS	11th	Learning time for sciences (minutes per week)
	DIRINS	13th	Teacher-directed instruction
	TMINS	14th	Learning time in total (minutes per week)
	PERFEED	18th	Perceived teacher feedback
	DISCLIMA	28th	Disciplinary climate in language classrooms (core classes in this PISA publication)
	BEINGBULLIED	29th	Experience of being bullied

Non-cognitive and metacognitive	BSMJ	2nd	Student's expected occupational status
	PISADIFF	7th	Perception of difficulty of the PISA test
	METASUM	15th	Metacognition: summarising
	STUBMI	19th	Student's body mass index
	GCSELFEFF	21st	Self-efficacy regarding global competence
	BELONG	22nd	Subjective well-being: sense of belonging at school
	SCREADDIFF	24th	Self-concept of reading: perception of difficulty
	GCAWARE	25th	Student awareness of global issues
	COGFLEX	26th	Student's cognitive flexibility/adaptability

Table 3 presents the variables that emerge as being the most important, with Stochastic Gradient Boosting, in secondary education: 8 relating to student background, 14 associated with school and 8 non-cognitive and metacognitive related to the student. In this stage, there are 29 from level 1 and only one from level 2: *student behaviour hindering learning* (ranked 11th). The most significant predictors here are revealed as: the *student's expected occupational status* (BSMJ), *learning time for sciences* (SMINS) and the index of economic, social and cultural status (ESCS). The descriptive analyses of the 30 most important variables shown in Table 2 and Table 3 can be found in the following link: https://acortar.link/hfDnvO (Tables S3 and S4).

TABLE 3. Most significant variables associated with grade repetition in secondary education.

Area	Label	Order	Variables
Student background	ESCS	3rd	Index of economic, social and cultural status
	CHANGE	5th	Number of changes in the educational biography
	HOMEPOS	9th	Home possessions
	BMMJ1	10th	Mother's ISEI
	HISEI	12th	Parents' highest occupational status
	HEDRES	20th	Home educational resources
	ICTRES	21st	Home ICT resources
	DURECEC	23rd	Duration of early childhood education and care

Related to school	SMINS	2nd	Learning time for sciences (minutes per week)
	LMINS	4th	Learning time for Language (minutes per week)
	MMINS	6th	Learning time for Mathematics (minutes per week)
	INFOCAR	7th	Information about careers
	STUBEHA	11th	Student behaviour hindering learning
	HOMESCH	13th	Use of ICT at home for school work
	TMINS	15th	Learning time in total (minutes per week)
	ICTOUTSIDE	17th	Subject-related use of ICT outside of school
	ENTUSE	18th	Use of ICT outside of school for leisure
	INFOJOB2	19th	Information about the labour market provided outside of school
	PERFEED	22nd	Perceived teacher feedback
	INFOJOB1	24th	Information about the labour market provided by the school
	DISCRIM	27th	Discriminatory school climate
	DIRINS	29th	Teacher-directed instruction
Non-cognitive and metacognitive	BSMJ	1st	Student's expected occupational status
	MASTGOAL	8th	Learning goals
	PISADIFF	14th	Perception of difficulty of the PISA test
	GCSELFEFF	16th	Self-efficacy regarding global competence
	RESILIENCE	25th	Student resilience
	METASUM	26th	Metacognition: summarising
	STUBMI	28th	Student's body mass index
	GCAWARE	30th	Student awareness of global issues

To achieve the second objective, once we had identified and sorted the variables with the strongest relationship to grade repetition, the next step was to define the predictive model, for which we dichotomised the categorical independent variables as indicated by Pardo and Ruiz (2013) (Table 4). For the predictor DURECEC, the OECD (2019) guidelines were followed.

TABLE 4. Recoding of independent variables for the multilevel logistic regression model.

Label	Categorical independent variable	Recoded values
SCCHANGE	Number of school changes	
CHANGE	Number of changes in the educational biography	0 = No change 1 = Once or more changes
IMMIG	Immigration background	0 = Immigrant 1 = Native
DURECEC	Duration of early childhood education and care	0 = Three or under 1 = Four or more

Table 5 summarises the regression models for grade repetition in primary education, with N1 or N2 after the variable name to identify the level to which it belongs. In model 1, all the student backgrounds are statistically significant, which explains the 62.380% variability in grade repetition. According to the odds ratio, the predictor *number of school changes* (SCCHANGE) presents the greatest probability of repetition. It should be noted that the *percentage of students from socioeconomically disadvantaged homes* (SC048Q03NA) is the only significant variable from level 2, and *mother's ISEI* (BMMJ1) has an odds ratio that is close to 1. Model 2, composed of the school-related variables, explains 19% of the criterion variance and, although they are all significant, four of them have an odds ratio of 1, with *experience of being bullied* (BEINGBULLIED) and *teacher-directed instruction* (DIRINS) being the main risk factors. Model 3 explains about 25% of the variance, containing only NC+MCC variables, of which *student awareness of*

global issues (GCAWARE) is not significant and has an odds ratio that is very close to the *body mass index* (STUBMI), with *perception of difficulty of the PISA test* (PISADIFF) being the variable with the greatest effect. The last model (M4) explains 61% of the variance in grade repetition in primary education; there is no significance attached to the effects of *learning time for sciences (minutes per week)* (SMINS), or the *body mass index* (STUBMI), and furthermore, there are eight variables with an odds ratio that is very close to 1; the variables with the most effect are *number of school changes* (SCCHANGE), *teacher-directed instruction* (DIRINS) and *perceived teacher feedback* (PERFEED).

Lastly, as Table 5 shows, regarding the model fit, model 4 has the lowest AIC and BIC score, with a significant reduction in the variance, equivalent to an R^2 of 9% compared to model 1. This is also significant in relation to the null model, with a reduction in the variance of model 1 (8%), model 2 (5%) and model 3 (9%).

TABLE 5. Odds ratio and standard errors for the multilevel logistic regression models in primary education.

Fixed effects	M0 (Null)	M1 (Backgrounds)	M2 (School)	M3 (NC+MCC)	M4 (All)
Intercept	0.10(0.03)***	0.14(0.08)***	0.07(0.09)***	0.41(0.11)***	0.38(0.17)***
ESCS_N1		0.59(0.03)***			0.67(0.03)***
SCCHANGE_N1		2.10(0.04)***			1.89(0.04)***
IMMIG_N1		0.62(0.05)***			0.58(0.05)***
ICTRES_N1		0.90(0.03)***			0.90(0.03)***
HEDRES_N1		0.89(0.02)***			0.93(0.03)***
DURECEC_N1		0.69(0.05)***			0.77(0.05)***
SC048Q03NA_N2		1.01(0.00)***			1.01(0.00)***
BMMJ1_N1		0.99(0.00)***			1.00(0.00)***
MMINS_N1			1.00(0.00)***		1.00(0.00)***
LMINS_N1			1.00(0.00)***		1.00(0.00)***
INFOCAR_N1			0.72(0.02)***		0.79(0.02)***
SMINS_N1			1.00(0.00)***		1.00(0.00)***
DIRINS_N1			1.30(0.03)***		1.24(0.02)***
TMINS_N1			1.00(0.00)***		1.00(0.00)***
PERFEED_N1			1.25(0.03)***		1.22(0.02)***
DISCLIMA_N1			0.83(0.02)***		0.89(0.02)***

BEINGBULLIED_N1			1.30(0.02)***		1.13(0.02)***
BSMJ_N1				0.97(0.00)***	0.98(0.00)***
PISADIFF_N1				1.31(0.02)***	1.19(0.02)***
METASUM_N1				0.69(0.02)***	0.75(0.02)***
STUBMI_N1				1.01(0.00)***	1.00(0.00)***
GCSELFEFF_N1				0.87(0.02)***	0.89(0.02)***
BELONG_N1				0.87(0.02)***	0.93(0.02)***
SCREADDIFF_N1				1.89(0.02)***	1.08(0.02)***
GCAWARE_N1				0.99(0.02)	1.01(0.03)***
COGFLEX_N1				1.08(0.02)***	1.06(0.02)***
Random effects					
Variance	0.52(0.05)***	0.20(0.03)***	0.42(0.04)***	0.39(0.04)***	0.20(0.07)***
PEV (%)		62.38	18.62	25.53	0.61
Fit indices					
AIC	22198.69	19804.01	21053.05	20120.57	18072.87
BIC	22215.58	19888.48	21145.96	20213.48	18309.368
Loglik	-11097.34	-9892.01	-10515.53	-10049.29	-9008.43
Deviance	20724.84	19026.15	19749.16	18906.55	17292.51
Chisq		2410.72	1163.65	2096.13	1767.1
Pr (> Chisq)		0.00	0.00	0.00	0.00

Note: *$p < .05$, **$p < .01$, ***$p < .001$. The reference category to interpret the odds ratio is the value 1 in all cases.

Lastly, Table 6 presents the regression models for grade repetition in secondary education, with N1 or N2 after the variable name to identify the level to which it belongs. Model 1 explains approximately 51% of the criterion and it is not significantly affected by *home ICT resources* (ICTRES), with the main risk factor being the *number of changes in educational biography* (CHANGE). Model 2 explains 32% of the variance; all the variables are statistically significant, although four of them have an odds ratio that is very close to 1. The greatest effect corresponds to *student behaviour hindering learning* (STUBEHA), a level 2 variable. Model 3, containing NC+MCC predictors, explains 14%. *Student resilience* (RESILIENCE) is not significant and there are two other predictors with an odds ratio that is very close to 1, with the greatest risk factor found in the variable *perception of difficulty of the PISA test* (PISADIFF). In the global model, *learning time for sciences* (SMINS), *learning time for mathematics* (MMINS) and *resilience* (RESILIENCE) are not significant and there are another six predictors with an odds ratio that is very close to 1. This set of predictors explains almost 56% of grade repetition in secondary education. The variables with the greatest effect are *number of school changes* (SCCHANGE), *student behaviour hindering learning* (STUBEHA) and *teacher-directed instruction* (DIRINS).

Lastly, it should be noted that model 4 provides the best fit, with a significant reduction in the variance, equivalent to an R^2 of 10.298% compared to model 1, although the reduction in model 1 is also significant in comparison to model 0 (7%), and in models 2 and 3 compared to the null model (5% and 10%, respectively).

TABLE 6. Odds ratio and standard errors for the multilevel logistic regression models in secondary education.

	M0 (Null)	M1 (Backgrounds)	M2 (School)	M3 (NC+MCC)	M4 (All)
Fixed effects					
Intercept	0.21(0.03)***	0.27(0.06)***	0.14(0.07)***	1.00(0.10)	0.70(0.13)***
ESCS_N1		0.67(0.02)***			0.76(0.02)***
CHANGE_N1		2.44(0.03)***			2.24(0.03)***
BMMJ1_N1		0.98(0.01)***			0.99(0.00)***
HEDRES_N1		0.85(0.02)***			0.92(0.02)***
ICTRES_N1		0.91(0.02)			0.91(0.02)***
DURECEC_N1		0.69(0.05)***			0.76(0.04)***
SMINS_N1			1.00(0.00)***		1.00(0.00)***
LMINS_N1			1.01(0.00)***		1.00(0.00)***

MMINS_N1			1.00(0.00)***		1.00(0.00)
INFOCAR_N1			0.72(0.02)***		0.76(0.02)***
STUBEHA_N2			1.42(0.03)***		1.25(0.02)***
HOMESCH_N1			0.91(0.02)***		0.93(0.02)***
TMINS_N1			1.00(0.00)***		1(0.00)***
ICTOUTSIDE_N1			0.89(0.02)***		0.96(0.02)***
ENTUSE_N1			1.10(0.02)***		1.13(0.02)***
INFOJOB2_N1			1.13(0.02)***		1.11(0.02)***
PERFEED_N1			1.16(0.02)***		1.15(0.02)***
INFOJOB1_N1			0.84(0.02)***		0.85(0.02)***
DISCRIM_N1			1.39(0.02)***		1.20(0.02)***
DIRINS_N1			1.21(0.02)***		1.21(0.02)***
BSMJ_N1				0.97(0.00)***	0.98(0.00)***
MASTGOAL_N1				0.78(0.02)***	0.82(0.02)***
PISADIFF_N1				1.23(0.02)***	1.14(0.02)***
GCSELFEFF_N1				0.86(0.02)***	0.88(0.02)***
RESILIENCE_N1				1.00(0.02)	0.99(0.02)
METASUM_N1				0.72(0.02)***	0.79(0.02)***
STUBMI_N1				1.01(0.00)***	1.01(0.00)***
GCAWARE_N1				1.08(0.02)***	0.76(0.02)***
Random effects					
Variance	0.70(0.04)***	0.34(0.05)***	0.47(0.04)***	0.60(0.05)***	0.32(0.04)***
PEV (%)		51	32.43	14	55.69
Fit indices					
AIC	32841,56	29736,77	30809,14	29611,70	26708,20
BIC	32858.46	29804.34	30944.28	29696.16	26961.58
Loglik	-16418.78	-14860.38	-15388.57	-14795.85	-13324.10
Deviance	30681.45	28238.71	29013.64	27662.13	25330.76
Chisq		3116.8	2060.3	152.57	6215.6
Pr (> Chisq)	0.70(0.04)***	0.34(0.05)***	0.47(0.04)***	0.60(0.05)***	0.32(0.04)***

Note: *p <.05, **p <.01, ***p <.001. The reference category to interpret the odds ratio is the value 1 in all cases.

4. Discussion

In the modelling produced, the most significant contextual indices have proved to be mainly from the first level, that is to say, student characteristics predominate over school characteristics. This result coincides with that obtained by Lopez et al. (2023) in their systematic review of research conducted with Spanish students. Also with other secondary analyses in which the object of study is performance, as in the study by Choi et al. (2018) with data on Spanish students from PISA 2012 and PIRLS 2006 (Progress in International Reading Literacy Study), or by Lopes et al. (2022) with data on Portuguese students from PIRLS 2016.

It can be concluded that the three variables that are most closely linked to grade repetition throughout compulsory education are *social status*, from the student background group; *learning time*, from the group of constructs associated with school; and the *level of expectation*, from the NC+MCC set, in a different order in the two stages. In this last group, the results are in line with those found by Asensio et al. (2018) and López et al. (2018), who, without differentiating the stages, establish expectations as one of the most important variables as regards performance. In terms of the importance of the student background variable, the present study is in line with the findings of Carabaña (2013), Choi et al. (2018) or García-Pérez et al. (2014). However, here we provide specific information relating to the differential weight depending on the stage, which may be interpreted as an indicator that the educational system is fulfilling its compensatory function, as

ESCS moves into third place in secondary education, behind a psychological variable and one related to the curriculum. Without a doubt, this evidence represents an interesting contribution by the present study. In the two stages, other significant psychoeducational elements emerge such as, from the set of school-related variables, *academic/professional information* (ranked tenth in primary education and seventh in secondary education), *teacher-directed instruction* (thirteenth place in primary education and twenty-ninth in secondary education), *perceived teacher feedback* and the *climate in the classroom*; and, from the NC+MCC group, *perception of the difficulty of the test*.

Specifically of note, in primary education only, is the relationship between grade repetition with *bullying* and *flexibility* or *cognitive adaptability* and in secondary education, with *use of ICT*, *learning goals* and *awareness of global issues*. Thus, this stage shows an increase in the number of school-related variables considered important, with a decrease in those related to background, on which the school has less power to act directly, thereby confirming and expanding the findings by Arroyo et al. (2019) with data from PISA 2015.

The emergence of differential models in both stages leads to the conclusion that how grade repetition is treated at the practical and political level needs to take this aspect into account. The existence of clear links between background variables and grade repetition also has implications for practice, as they are a warning of the need to persevere with compensatory ap-

proaches that improve equity, particularly in primary education. Furthermore, the importance of the time spent on learning indicates practices aimed at efficient management of this variable, without being restricted to simply increasing study time in disadvantaged classes, which is a problem raised, for example, by the policies of summer courses to recover "educational debts" (Battistin & Schizzerotto, 2019). The most equitable solution would be to support the principles of individualisation, early intervention (Choi et al., 2018) and inclusion (Tapia & Álvarez, 2022), in which teacher training and collaborative teaching or team teaching are portrayed as promising strategies in the most successful educational systems (Niemi, 2015). In terms of the level of expectation, it is important to work on motivation, at both the primary and secondary stages (Rhodes et al., 2018; Rodríguez-Rodríguez, 2022). Other specific aspects that should be developed in order to lessen the rate of grade repetition in both stages are feedback on learning, the classroom climate or academic/professional guidance.

The availability of such an extensive database as PISA has enabled the use of machine learning algorithms in this work, which is one of its main contributions. The multilevel study conducted was based on a prior selection that helped to minimise specification error, which is difficult to avoid in educational research. The use of machine learning techniques complements and increases the sensitivity of hierarchical regression models, by providing empirical criteria to maintain the important predictors and dismiss the ones that are irrelevant. Although it has not proved to be so efficient as regards the latter function due to the existence of significant odds ratios, albeit equal or very close to 1, this is a methodology that allows for a quantitative calculation not only of each predictor's worth but also, fundamentally, the value of the model's predictive validity.

In terms of limitations, it should be noted that machine learning methods may overfit models to the current sample's characteristics. Additionally, the PISA background questionnaires are based on self-reporting, which is not free of measurement errors. Moreover, it is worth mentioning that the relationships identified do not imply causality, so for this reason the results of this type of exploratory study should be complemented by confirmatory research posing hypotheses on the direct and indirect effects that apply to the factors involved. In this sense, any educational measure that is adopted to reduce grade repetition, given the relationship of this variable with performance, will have an indirect effect on it, although this aspect should be studied specifically using covariance structure modelling, or designs that are experimental or quasi-experimental. For the future, it would also be advisable to examine the replicability of the results in depth, by comparing the results in participant subgroups, for example, by gender. Likewise, it would be interesting to explore the factors associated with grade repetition by extending the hierarchy levels to include teachers (three-level model) and even autonomous regions (four-level model).

Authors' contributions

Alexander Constante-Amores: Data curation; Formal analysis; Methodology; Writing (original draft).

Delia Arroyo-Resino: Conceptualisation; Formal analysis; Writing (original draft); Writing (review and editing).

María Sánchez-Munilla: Data curation; Formal analysis; Methodology; Writing (original draft).

Inmaculada Asensio-Muñoz: Conceptualisation; Supervision; Writing (original draft); Writing (review and editing).

References

Allen, C. S., Chen, Q., Willson, V. L., & Hughes J. N. (2009). Quality of design moderates effects of grade retention on achievement: A metaanalytic, multi-level analysis. *Educational Evaluation and Policy Analysis*, *31* (4), 480-99. https://doi.org/10.3102/0162373709352239

Arroyo, D., Constante-Amores, A., & Asensio, I. (2019). La repetición de curso a debate: un estudio empírico a partir de PISA 2015 [Debate on grade repetition: An empirical study from PISA 2015]. *Educación XX1, 22* (2), 69-92. https://doi.org/10.5944/educxx1.22479

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

Arroyo, D., Constante-Amores, A., Gil, P., & Carrillo, P. J. (2024b). Student well-being and mathematical literacy performance in PISA 2018: A machine-learning approach. *Educational Psychology*, *44* (3), 340-357. https://doi.org/10.1080/01443410.2024.2359104

Asensio, I., Carpintero, E., Expósito, E., & López, E. (2018). ¿Cuánto oro hay entre la arena? Minería de datos con los resultados de España en PISA 2015 [How much gold is in the sand? Data mining with Spain's PISA 2015 results]. **Revista Española de Pedagogía**, *76* (270), 225-245. https://doi.org/10.22550/REP76-2-2018-02

Battistin, E., & Schizzerotto, A. (2019). Threat of grade retention, remedial education and student achievement: Evidence from upper secondary schools in Italy. *Empirical Economics*, *56*, 651-678. https://doi.org/10.1007/s00181-018-1443-6

Cameron, A. C., & Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, *77* (2), 329-342. https://doi.org/10.1016/S0304-4076(96)01818-0

Carabaña, J. (2013). Repetición de curso y puntuaciones PISA, ¿cuál causa cuál? [Grade repetition and PISA scores, which causes which?] In Instituto Nacional de Evaluación Educativa (Ed.), *PISA 2012: Programa para la evaluación de los alumnos, informe español (volumen II: análisis secundario) [PISA 2012: Programme for international student assessment, Spanish report (volume II: secondary analysis)]*. INEE.

Choi, A., Gil, M., Mediavilla, M., & Valbuena, J. (2018). Predictors and effects of grade repetition. *Revista de Economía Mundial*, (48), 21-42.

Constante, A., Florenciano, E., Navarro, E., & Fernández, M. (2021). Factores asociados al abandono universitario [Factors associated with university dropout]. *Educación XX1, 24* (1), 17-44. https://doi.org/10.5944/educxx1.26889

Cordero, J., Manchón, C., & Simancas, R. (2014). La repetición de curso y sus factores condicionantes en España [Repetition and explanatory factors in Spain]. *Revista de Educación*, (365), 12-37. https://doi.org/10.4438/1988-592X-RE-2014-365-263

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self- concept: Reconsidering structural relationships. *Learning and Individual Differences, 19* (4), 499-505. https://doi.org/10.1016/j.lindif.2009.05.004

Fernández-Lasarte, O., Ramos-Díaz, E., & Axpe, I. (2019). Rendimiento académico, apoyo social percibido e inteligencia emocional en la universidad [Academic performance, perceived social support and emotional intelligence at the university]. *European Journal of Investigation in Health*, *9* (1), 39-49.

Gamazo, A., & Martínez-Abad, F. (2020). An exploration of factors linked to academic performance in PISA 2018 through data mining techniques. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.575167

García-Pérez, J., Hidalgo, M., & Robles, J. (2014). Does grade retention affect students' achievement? Some evidence from Spain. *Applied Economics*, *46* (12), 1373-1392. https://doi.org/10.1080/00036846.2013.872761

Gaviria, J. L., & Castro, M. (2005). *Modelos jerárquicos lineales [Linear hierarchical models]*. La Muralla.

González-Nuevo, C., Postigo, Á., García-Cueto, E. Menéndez-Aller, Á., Muñiz, J., Cuesta, M., Álvarez-Díaz, M., & Fernández-Alonso, R. (2023). Grade retention impact on academic self-concept: a longitudinal perspective. *School Mental Health, 15*, 600-610. https://doi.org/10.1007/s12310-023-09573-2

Gorostiaga, A., & Rojo-Álvarez, J. L. (2016). On the use of conventional and statistical-learning techniques for the analysis of PISA results in Spain. *Neurocomputing*, *171*, 625-637. https://doi.org/10.1016/j.neucom.2015.07.001

Hattie, J. (2017). *Aprendizaje visible para profesores. Maximizando el impacto en el aprendizaje [Visible learning for teachers. Maximising learning impact]*. Paraninfo.

Hornstra, L., Majoor, M., & Peetsma, T. (2017). Achievement goal profiles and developments in effort and achievement in upper elementary school. *British Journal of Educational Psychology, 87* (4), 606-629. https://doi.org/10.1111/bjep.12167

Ikeda, M., & García, E. (2014). Grade repetition: A comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies*, (1), 269-315. https://doi.org/10.1787/eco_studies-2013-5k3w65mx3hnx

Jimerson, S. R. (2001). Meta-analysis of grade retention research: new directions for practice in the 21st century. *School Psychology Review*, *30* (3), 420-437. http://dx.doi.org/10.1080/02796015.2001.12086124

Jimerson, S. R., & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School psychology quarterly*, *22* (3), 314.

Kılıç, D., Aşkın, Ö. E., & Öz, E. (2017). Identifying the classification performances of educational data mining methods: A case study for TIMSS. *Kuram ve Uygulamada Egitim Bilimleri*, *17* (5), 1605-1623. https://jestp.com/menuscript/index.php/estp/article/view/426

Kiray, S. A., Gok, B., & Bozkir, A. S. (2015). Identifying the factors affecting science and mathematics achievement using data mining methods. *Journal of Education in Science, Environment and Health*, *1* (1), 28-48.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, *35* (2), 125-141. https://doi.org/10.1207/S15326985EP3502_6

Liu, X., & Ruiz, M. E. (2008). Using data mining to predict K-12 students' performance on large-scale assessment items related to energy. *Journal of Research in Science Teaching*, *45* (5), 554-573. https://doi.org/10.1002/tea.20232

Lopes, J., Oliveira, C. & Costa, P. (2022). Determinantes escolares y de los estudiantes en el rendimiento lector: un análisis multinivel con estudiantes portugueses. *Revista de Psicodidáctica*, *27*, 29-37. http://doi.org/10.1016/j.psicod.2021.05.001

López, E., Expósito, E., Carpintero, E., & Asensio, I. (2018). ¿Qué nos dice PISA sobre la enseñanza y el aprendizaje de las ciencias? Una aproximación a través de árboles de decisión [What does PISA tell us about science teaching and learning? A decision tree approach]. *Revista de Educación*, (382), 133-161.

López, F., & García, I. (2020). *A vueltas con la equidad en educación: una aproximación empírica en la perspectiva de las consecuencias [The equity in education: An empirical approach in the perspective of consequences]*. Universidad Camilo José Cela. http://hdl.handle.net/20.500.12020/901

López, L., González-Rodríguez, D., & Vieira, M. J. (2023). Variables que afectan la repetición en la educación obligatoria en España [Variables affecting repetition in compulsory education in Spain]. *Revista Electrónica de Investigación Educativa*, *25*, 1-15. https://doi.org/10.24320/redie.2023.25.e17.4942

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España. Análisis empíricos y recomendaciones políticas [Grade repetition and graduation in compulsory secondary education in Spain. Empirical analysis and policy recommendations]. *Revista de Educación*, (394), 325-353. https://doi.org/10.4438/1988-592X-RE-2021-394-510

Marsh, H. W., Parker, P. D., & Pekrun, R. (2018). Three paradoxical effects on academic self-concept across countries, schools, and students. *European Psychologist*, *24* (3), 231-242. http://doi.org/10.1027/1016-9040/a000332

Mathys, C., Véronneau, M., & Lecocq, A. (2019). Grade retention at the transition to secondary school: Using propensity score matching to identify consequences on psychosocial adjustment. *Journal of Early Adolescence*, *39* (1), 97-133. https://doi.org/10.1177/0272431617735651

Martínez-Abad, F., & Chaparro, A. A. (2017). Data-mining techniques in detecting factors linked to academic achievement. *School Effectiveness and School Improvement*, *28* (1), 39-55. https://doi.org/10.1080/09243453.2016.1235591

Martínez-Abad, F., Gamazo, A., & Rodríguez-Conde, M. J. (2020). Educational data mining: Identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation*, *66*, 100875. https://doi.org/10.1016/j.stueduc.2020.100875

Nieto-Isidro, S., & Martínez-Abad, F. (2023). Repetición de curso y su relación con variables socioeconómicas y educativas en España [Grade retention and its relationship with socioeconomic and educative variables in Spain]. *Revista de Educación*, *1* (402), 207-236. https://doi.org/10.4438/1988-592X-RE-2023-402-600

Ministerio de Educación y Formación Profesional. (2019). *PISA 2018. Programa para la evaluación internacional de los estudiantes. Informe español (vol. I) [PISA 2018. Programme for international student assessment. Spanish report (vol. I)].* https://www.educacionyfp.gob.es/inee/evaluaciones-internacionales/pisa/pisa-2018/pisa-2018-informes-es.html

Ministerio de Educación y Formación Profesional. (2022). *Sistema estatal de indicadores de la educación 2022 [State system of education indicators 2022].* https://www.libreria.educacion.gob.es/libro/sistema-estatal-de-indicadores-de-la-educacion-2022_184171/

Niemi, H. (2015). Teacher professional development in Finland: Towards a more holistic approach. *Psychology, Society and Education*, 7 (3), 279-294.

OCDE (2019). *PISA 2018 assessment and analytical framework*. https://doi.org/10.1787/b25efab8-en

OCDE (2020). *PISA 2018 results (volume VI). Are students ready to thrive in an interconnected world?* https://www.oecd.org/publications/pisa-2018-results-volume-vi-d5f68679-en.htm

Pardo, M., & Ruiz, M. (2013). *Análisis de datos en ciencias sociales y de la salud III [Data analysis in health and social sciences III].* Síntesis.

Peixoto, F., Monteiro, V., Mata, L., Sanches, C., Pipa, J., & Navas, L. (2016). *To be or not to be retained… That's the question!* Retention, self-esteem, self-concept, achievement goals, and grades. *Frontiers in Psychology*, 7, (1550). https://doi.org/10.3389/fpsyg.2016.01550

Raschka, S. (2015). *Python machine learning*. Packt.

Rhodes, J., Thomas, J. M., & Liles, A. R. (2018). Predictors of grade retention among children in an elementary school truancy intervention. *Journal of At-Risk Issues*, *21* (1), 1-10.

Rivero, G. (2011*). Análisis de datos incompletos en ciencias sociales [Incomplete data analysis in social sciences].* CIS.

Rodríguez-Rodríguez, D. (2022). Repetición de curso, rendimiento académico y variables motivacionales en Educación Secundaria Obligatoria: un estudio longitudinal [Grade retention, academic performance and motivational variables in compulsory secondary education: A longitudinal study]. *Psicothema*, *34* (3), 429-436. https://doi.org/10.7334/psicothema2021.582

Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis*. McGraw-Hill

Sarkar, D., Bali, R., & Sharma, T. (2018). *Practical machine learning with Python. A problem-solvers guide to building real-world intelligent systems*. Apress.

Schwerdt, G., West, M. R., & Winters, M. A. (2017). The effects of testbased retention on student outcomes over time: Regression discontinuity evidence from Florida. *Journal of Public Economics*, *152*, 154-169. https://doi.org/10.1016/j.jpubeco.2017.06.004

Seabra, D. A., & Ferrão, M. E. (2016). Repetência e indisciplina: Evidências de Brasil e Portugal no PISA 2012 [Grade repetition and indiscipline: Evidence from Brazil and Portugal in Pisa, 2012]. *Cadernos de Pesquisa*, *46* (161), 614-636. https://publicacoes.fcc.org.br/cp/article/view/3669

Tapia, J. G., & Álvarez, C. A. (2022). La repetición de curso en educación secundaria en clave organizacional [Grade repetition in secondary education in an organizational perspective]. *Avances en Supervisión Educativa*, (38). https://doi.org/10.23824/ase.v0i38.772

Urbina, A. B., & De la Calleja, J. (2017). Breve revisión de aplicaciones educativas utilizando minería de datos y aprendizaje automático [Brief review of educational applications using data mining and machine learning]. *Revista Electrónica de Investigación Educativa*, *19* (4), 84-96. https://doi.org/10.24320/redie.2017.19.4.1305

Valbuena, J., Mediavilla, M., Choi, Á., & Gil, M. (2021). Effects of grade retention policies: A literature review of empirical studies applying causal inference. *Journal of Economic Surveys*, *35* (2), 408-451. https://doi.org/10.1111/joes.12406

Van Canegem, T., Van Houtte, M., & Demanet, J. (2021). Grade retention and academic self-concept: A multilevel analysis of the effects of schools' retention composition. *British Educational Research Journal*, *47* (5), 1340-1360. https://doi.org/10.1002/berj.3729

Warren, J. R., Hoffman, E., & Andrew, M. (2014). Patterns and trends in grade retention rates in the United States, 1995-2010. *Educational Researcher*, *43* (9), 433-443. https://dx.doi.org/10.3102/0013189X14563599

Authors' biographies

Alexander Constante-Amore. Bachelor's Degree in Pedagogy and Master's Degree in Educational Research from the Universidad Complutense de Madrid (UCM). He is currently completing the Doctoral Programme in Education at the UCM. He is a lecturer at the Universidad Camilo José Cela and the Universidad Europea de Madrid, where he teaches subjects in the areas of Statistics, Biostatistics and Research Methods. He was an intern at the Instituto Nacional de Estadística [Spanish National Statistics Institute] and he is a member of the research group Medida y Evaluación de Sistemas Educativos (MESE) [Measurement and Evaluation of Educational Systems]. His line of research focuses on the evaluation of educational systems.

https://orcid.org/0000-0003-2024-9615

Delia Arroyo-Resino. Lecturer at the Universidad Complutense de Madrid. She holds an International Doctorate with an Outstanding Award from the Department of Research and Diagnostic Methods in Education for the Universidad Complutense de Madrid. She is a member of the research group Medida y Evaluación de Sistemas Educativos (MESE).

https://orcid.org/0000-0002-3784-7745

María Sánchez-Munilla. Bachelor's Degree in Pedagogy and a Master's Degree in Methodologies in Behavioural Health Sciences at the Universidad Nacional de Educación a Distancia (UNED). She is currently completing the UCM Doctoral Programme in Education on a predoctoral contract named Formación del Profesorado Universitario (FPU) [University Teacher Training]. She is a member of the research group Medida y Evaluación

de Sistemas Educativos (MESE) and the Servicio de Evaluación y Diagnóstico en Educación (SEDE) [Educational Evaluation and Diagnosis Service] of the Faculty of Education at the UCM

https://orcid.org/0000-0001-6689-7170

Inmaculada Asensio-Muñoz. Doctorate with an outstanding award from the Universidad Complutense de Madrid. Senior lecturer in the area of Research and Diagnostic Methods in Education, at the Department of Research and Psychology in Education in the Faculty of Education at this University. She has extensive teaching experience in pedagogy and teacher training, both at undergraduate and master's or doctorate level. She is a member of the research group Medida y Evaluación de Sistemas Educativos (MESE) and, as an expert in educational research methodology, during her professional career, she has worked on several funded projects on research and innovation related to this subject. Her publications focus on improving education and learning, and, in general, on the role of teaching and guidance.

https://orcid.org/0000-0003-3523-570X

A full-time professor dedicated to you: ChatGPT from university students' perspective

Un profesor particular a tiempo completo: ChatGPT desde el punto de vista de los estudiantes universitarios

Dragana PAVLOVIC, PhD. Associate Professor. University of Niš, Serbia (*dragana.pavlovic@filfak.ni.ac.rs*).

Joan SOLER-ADILLON, PhD. Associate Professor. Universitat Oberta de Catalunya (UOC), Spain (*jsoleradillon@uoc.edu*).

Zorica STANISAVLJEVIC-PETROVIC, PhD. Full Professor. University of Niš, Serbia (*zorica.stanisavljevic.petrovic@filfak.ni.ac.rs*).

Abstract:

Over the last decades, the emergence of new technologies has brought changes in all areas of social life, including the education sector. Among those, artificial intelligence (AI), after many years playing a secondary role, has recently positioned itself as one with the most capacity for disruption. The newest AI tools have the potential to meaningfully transform various education areas, particularly at the university level. This paper presents a comparative analysis of the opinions of media students from Spain and Serbia on the use of ChatGPT in the learning process. The aim of the work is to determine the similarities and differences in the opinions of students in relation to the following aspects: intention to use, benefits and challenges and potential for personalised learning. A qualitative research method, with structured written interview, was used on a sample of 40 students of the initial years of media studies from universities in both countries. The results indicate that there are similarities and differences in the opinions of media students about using ChatGPT for learning. The findings provide insight into the opinions of media students as key subjects in the learning process and can be a starting point for further research aimed at understanding the role of AI tools and their application in higher education.

Keywords: artificial intelligence, ChatGPT, university, Spain, Serbia, personalised learning.

Resumen:

En las últimas décadas, la aparición de nuevas tecnologías ha traído cambios en

todas las áreas de la vida social, incluido el sector educativo. Entre ellas, la inteligencia artificial (IA), tras haber desempeñado durante años un papel secundario, se ha posicionado recientemente como una de las de mayor capacidad de disrupción. Las últimas herramientas de IA tienen el potencial de transformar de forma significativa varias áreas de la educación, en especial en el nivel universitario. Este artículo presenta un análisis comparativo de las opiniones de estudiantes universitarios en España y Serbia sobre el uso de ChatGPT en el proceso de aprendizaje. El objetivo del trabajo es determinar las similitudes y diferencias entre sus opiniones con relación a intención de uso, ventajas e inconvenientes, y potencial para el aprendizaje personalizado. La investigación se ha basado en el método de investigación

cualitativa, con una entrevista estructurada por escrito. La muestra ha estado compuesta por 40 estudiantes de universidades de ambos países, todos ellos matriculados en primero o segundo año de carreras relacionadas con los medios y el multimedia. Los resultados apuntan a similitudes y diferencias en las opiniones respecto al uso de ChatGPT. Así, proporcionan una panorámica de la visión del alumnado como sujetos clave en el proceso de aprendizaje y pueden ser un punto de partida para investigaciones futuras destinadas a comprender el papel que deben desempeñar las herramientas de IA y su aplicación en la educación superior.

Palabras clave: inteligencia artificial, ChatGPT, universidad, España, Serbia, educación personalizada.

1. Introduction

Over recent decades, the rise of new technologies, especially artificial intelligence (AI), has significantly impacted various sectors, including education. AI, defined as machines programmed to simulate human intelligence (Russell & Norvig, 2010), is reshaping educational practices across all levels, notably in higher education (Klutka et al., 2020; Kuleto et al., 2021). In university settings, AI's influence is profound across four key areas: profiling and prediction, assessment and evaluation, adaptive systems, and intelligent teaching systems, proving increasingly effective in academic environments (Kim et al., 2020; Woo & Choi, 2021; García-Peñalvo,

2023; Farrokhnia et al., 2023). This technological shift is particularly relevant for media studies, a field expected to be revolutionized by AI (Pavlik, 2023; Dhiman, 2022).

The current debate on AI in education includes its current impact on pedagogical, teaching, and cognitive processes, with ongoing concerns about ethical issues like transparency and credibility (Bozkurt, 2023; Dwivedi et al., 2023; Kasneci et al., 2023; Holmes et al., 2023). Research emphasizes AI's potential to transform both formal and informal educational settings, moving from AI-directed to AI-supported and AI-empowered frameworks, where

learners progressively lead their educational journeys (Schiff, 2021; Ouyang & Jiao, 2021; Adiguzel et al., 2023).

There's a growing need for studies exploring how students adapt to AI-enhanced learning environments while considering the diversity of influencing factors. Recent studies have started to explore emotional responses of students to AI in education, suggesting future applications might offer even greater innovation (Ezquerra et al., 2023; Ezquerra et al., 2022). This paper adds to the discourse by examining student perspectives from two countries on AI's role in their educational experiences.

2. Theoretical approach

In 2022, the educational landscape was notably affected by the advent of ChatGPT, an AI tool developed by OpenAI, which quickly gained widespread popularity (Haleem et al., 2022). Within two months of its release in November 2022, ChatGPT had attracted 100 million users (Sabzalieva & Valentini, 2023), illustrating its broad impact and potential in various educational domains, including programming, natural sciences, medicine, social sciences (Tian et al., 2023; Wardat et al., 2023; Lee, 2023a; Tiunova & Muñoz, 2023).

ChatGPT, primarily functioning as a language model, supports educational activities by facilitating research, analysis, and various forms of writing (Raman et al., 2023). Its capacity to understand complex sentence structures and gener-

ate high-quality responses has made it a very valuable tool in academia. It assists in generating ideas for research, analysis, and various writing tasks, including essays, term papers, and articles (Rasul et al., 2023; Fitria, 2023; Avila-Chauvet & Mejía, 2023). It also provides direct interaction with students, offering rapid answers across a range of subjects (Hariri, 2023; Opara et al., 2023), and is instrumental in tasks such as proofreading and translating (Jiao et al., 2023).

In higher education, ChatGPT's benefits extend to enhancing administrative services and improving university operations. It supports adaptive learning, provides personalized feedback, and assists in developing innovative assessment methods. Its accessibility, being free and available to a broad user base, allows for the creation of flexible virtual educational environments conducive to learning from any location at any time (Rasul et al., 2023; Halaweh, 2023; Yu, 2023). Sok and Heng (2023) emphasize its role in innovating and enhancing pedagogical practices, developing learning assessments, providing virtual personal tutoring, assisting with draft creation, and facilitating idea generation.

The tool is particularly praised for its ability to facilitate personalized learning tailored to the diverse needs of students. This includes creating customized exercises, providing feedback, and developing educational materials that reflect individual student progress (Atlas, 2023; Baidoo-Anu & Ansah, 2023; Woolf et al., 2023; Hong, 2023).

Moreover, it fosters a collaborative learning environment, enhancing critical thinking and facilitating knowledge transfer across various contexts (Rodrigues & dos Santos, 2023; Sison et al., 2023; Mollick & Mollick, 2022). Studies indicate that ChatGPT aids in developing learning skills by providing feedback that succinctly and coherently evaluates student performance (Dai et al., 2023; Firat, 2023). Furthermore, ChatGPT is seen as beneficial in the educational process, significantly enhancing academic outcomes through its capacity to process extensive data and learn from user interactions (Alshater, 2023; Dergaa et al., 2023; Romero-Rodríguez et al., 2023).

However, the deployment of ChatGPT also presents challenges, particularly concerning academic integrity, reliability, and potential perpetuation of biases (Rasul et al., 2023). Qadir (2023) notes that generative systems like ChatGPT may perpetuate biases and disseminate incorrect or incomplete information, rooted in the limitations of their data sources. Additionally, the temporally restricted knowledge base of ChatGPT (particularly in its free version, which is limited to data up to September 2021) may be a disadvantage for students working on less known, highly specialized, or very recent topics (Megahed et al., 2023; Rudolph et al., 2023; Sabzalieva & Valentini, 2023).

The issue of information credibility with AI systems like ChatGPT often creates mistrust and doubts regarding the accuracy and reliability of the data they produce, potentially undermining student confidence. Health (2023) reports that the developers of ChatGPT acknowledge its potential for generating inaccurate or biased results, and its occasional reference to non-existent articles or support of prejudices. Rose (2023) critiques ChatGPT for bias, inappropriate responses, and a lack of emotional intelligence. Uludag (2023) highlights its deficiencies in responding to emotional cues, understanding context, generating original ideas and creative solutions. Additionally, from a psychological perspective, complaints about ChatGPT include its lack of empathy and inability to provide adequate support and understanding to users (Kalla & Smith, 2023; Biswas, 2023).

The deployment of ChatGPT raises significant ethical issues and legal risks, including environmental impacts, content moderation challenges, and copyright infringement risks (Baskara, 2023; Iskender, 2023; Sullivan et al., 2023). Additionally, there are legal concerns related to privacy violations and data security (Addington, 2023; Sebastian, 2023), and the potential for academic plagiarism and other breaches of intellectual property, which pose threats to academic integrity (Susnjak, 2022; Kitamura, 2023; Frye, 2022). In this context, many scholars express deep ethical concerns about severe negative outcomes, such as cheating, dishonesty, deception, or manipulation, which could detrimentally impact both educators and students

(Tlili et al., 2023; Cotton et al., 2023; Zhuo et al., 2023).

To address the challenges associated with ChatGPT, it is key to enhance digital literacy and foster a critical perspective among students regarding the capabilities and limitations of AI technologies. Educators play a pivotal role in shaping realistic perceptions of AI tools by integrating them thoughtfully into curricula and pedagogical practices. Additionally, the use of ChatGPT in higher education necessitates the development of innovative assessment methods that emphasize creativity and critical thinking, skills that AI has yet to master (Zhai, 2023).

3. Methodological approach

This paper presents a comparative analysis of the opinions of media students from Spain and Serbia about the potential of ChatGPT in the learning process. In accordance with the set goal, the research tasks are aimed at examining the similarities and differences in the opinions of students from both countries on the use of ChatGPT in the following aspects: (1) intentions and purposes of its use, (2) benefits and challenges and (3) potential for personalised learning. According to the nature of the research problem, we applied a descriptive research design (Eze et al., 2018), with the use of qualitative and comparative research methods. The basic instrument of the research was a written structured interview, with questions related to various aspects of the application of this AI tool in learning. Structured interviews

were chosen to ensure consistency in the data collection process across two diverse cultural and educational contexts: Spain and Serbia. Previous research has shown that this method enables a standardised and objective approach to data collection, i.e. reducing the researcher's bias (see Lee, 2003; Silva, 2007; Kallio et al., 2016).

The interview questions enquire about the student's subjective opinion about ChatGPT in the learning process. The overall duration of the interview was around 45 minutes. The survey was conducted during the first half of May 2023. The research sample is deliberate and uniform and consists of a total of 40 media studies first and second year students: 20 from Spain (Universitat de Vic - Universitat Central de Catalunya) and 20 from Serbia (University of Niš). Students from Spain study Multimedia (Applications and Videogames). This program trains professionals specializing in digital creation: concept generation, graphic design, development and programming, production, and post-production. Students from Serbia are from the Communication and Journalism Department, Faculty of Philosophy, and they study a variety of multidisciplinary subjects related to multimedia applications (computer literacy, modern media technologies, new media, etc.). The choice of first- and second-year media studies students as research participants was influenced by the consideration that these students have a fresh perspective, greater openness to new technologies and can provide insight into the long-

term impact of technology on their educational path.

The procedure included the process of checking and coding the completed instruments, whereby the interviews filled in by students from Spain were marked with the initial letter of the country, E, and a serial number from 1-20 (E1 to E20), while those of students from Serbia were marked with the initial letter of the country, S, and serial number from 1-20 (S1-S20). Research data were processed qualitatively in accordance with established research methodologies (Mathers et al., 2002), and presented graphically. For processing, the thematic analysis was used to interpret the responses from the interviews. This involved a detailed coding process where responses were initially read to identify key themes that had emerged consistently across the dataset. These themes were aligned with our research objectives, focusing on the students' intentions, benefits perceived, challenges faced, and the potential for personalized learning using ChatGPT.

The descriptive research design was selected because it is well-suited for studies that aim to describe the characteristics of a specific group or phenomena. On writing the final version of this paper, ChatGPT was briefly used to get suggestions on how to shorten the text in order to fulfill the final publication requirements (OpenAI, 2023).

In the academic literature, Spain and Serbia are frequently compared, particu-

larly in the context of higher education and general educational issues. This comparative analysis spans various studies and includes multiple aspects of education (Kolenc, 2011; Fedorov & Levitskaya, 2015; Despotovic et al., 2019; Dolenec et al., 2020; Brkanlić et al., 2020; Corbí et al., 2021; Podstawski et al., 2022; Borsos et al., 2022).

4. Results

4.1. Intentions and purposes for using ChatGPT for learning

In the initial part of the interview, students were asked about their intentions and usage of ChatGPT in their learning and study processes. The responses indicated that a substantial majority, 80%, of students from both Spain and Serbia have utilized ChatGPT for educational purposes. Among these respondents, only one Spanish student was not previously aware of ChatGPT. Conversely, a smaller segment of Serbian students, 20%, experimented with specific features of the tool but did not employ it extensively in their learning.

Students from both countries who did use ChatGPT for educational purposes gave very uniform answers. Among the most repeated answers are those related to quickly finding information, writing essays, preparing exams and homework, making presentations, literature search, plot writing, help defining guidelines for specific topics and for translation purposes. In the answers of students from Spain, the main purpose of using

ChatGPT in learning is the creation of homework (E7); for practical works and summaries (E11) or for checking certain information: "I used Chat GPT during the process of studying to ask [about] certain doubts" (E14). Students from Serbia, similarly, used this tool for exam preparation: "I used it several times during exam preparation, and I used it several times to search literature for seminar papers" (S14), to search for answers to some exam questions or for guidance for writing on a certain topic (S10).

Intention to use ChatGPT in further studies is shown in Figure 1. Students from Spain and Serbia have the most uniform opinion regarding the future use of ChatGPT for organising and creating presentations, with 55% and 65% of positive answers respectively. Students from both countries cited some advantages that the AI tool offers for this particular task, among which the structuring of the presentation, as well as the clarity and precision of the text, stand out. In support of this, they state that "it is a good option if you have problems structuring a presentation. Not a real fan on that aspect, but I can see myself using it in the future" (E15) and also that "yes, it would make presentations easier, because the text generated by ChatGPT is concise and clear" (S10). In contrast, students who declared that they have no intention of using this function explain their attitude as follows: "I don't think so. I need to already understand well the material to be able to know if the examples of ChatGPT are correct" (E3); "I'm not planning for now because I don't think he's giving complete instructions, and I can't rely on him to come up with a whole

FIGURE 1. Students' intentions and purposes for using ChatGPT.
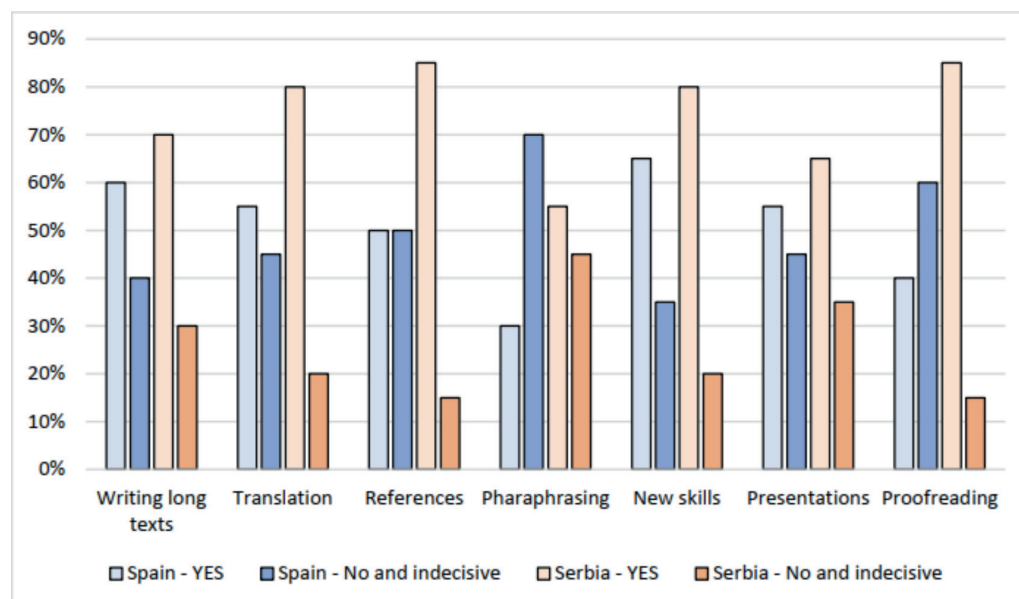
presentation plan for me. I'll better figure out what and how to show it myself; I can possibly look for examples if I don't have inspiration" (S4).

On the questions related to writing long texts (essays, term papers, project proposals, articles and others), 60% of students from Spain and 70% of students from Serbia expressed their intention to use this feature. According to them, ChatGPT can be useful in the production of ideas and motivation for writing; for example, "for some research proj-ects, also some memory projects. Sometimes to look up programming code. Also, I will use it to brainstorm ideas for some projects" (E16). Similarly, students from Serbia state that they plan to use it "as a starting point for making a plan and sequence; I would also use it to collect some information, although I would never completely surrender and reduce everything to ChatGPT" (S1). Likewise, they point out that, thanks to ChatGPT, "that boring beginning of every paper writing is made easier, since it provides guidelines, a theoretical basis and sources of information that I can research on my own. In this sense, he is an excellent motivator because he provides basic information at the start, which can certainly be useful for studying and writing student assignments" (S19).

Students who do not intend to use Chat GPT to write long texts often express doubts about its reliability and relevance of information in their explanations, but they think that it can be good

for providing initial ideas. In support of this, some stated that they do not plan to use it because they want to preserve the authenticity of their writing style: "I do like to write my way, but it's really good to give ideas about topic or to structure a really long text" (E15); "I don't plan to use it to write student papers, because I'm proud of my papers! But I will definitely use it to find other works, relevant information that I will later use in writing and studying" (S17). Some see ChatGPT as a distraction: "After a few months of using ChatGPT, I found that it has the advantage of formulating answers in a human way, but it doesn't necessarily make it useful. I think I will refer to it for simple short questions, but in terms of big projects I argue that it is more likely to distract me than to be helpful" (E8).

The biggest difference in students' answers about the intention of use can be observed regarding proofreading (checking grammar, spelling and style) of written text. As many as 85% of students from Serbia expressed their intention to use ChatGPT for this in the future, compared to only 40% of students from Spain. The former group sees ChatGPT as an accessible tool that "exactly states what we did wrong, which grammatical norm we violated" (S1) and use it "because I'm often not sure if everything is correct and I want everything to be good and proper" (S10). Also, a student from Serbia states that "sometimes I have a block while writing a text and compose quite simple sentences, and ChatGPT helps me to improve my style in order to explain

my thoughts" (S14). Students from Spain who intend to use this tool for proofreading explain it in the following way: "Why not? A computer can outperform a human in a logic-based context 9 times out of 10. Furthermore, what would be the harm?" (E20). However, those students who do not see the potential for proofreading state: "No, I have online correctors for that" (E19); "No I don't. I usually write in Catalan and there are better tools for that" (E15).

4.2. Benefits and challenges of ChatGPT in learning

Figure 2 shows the answers of the second part of the interview, intended to determine the opinions of students regarding the benefits and challenges of using ChatGPT in the learning process. With a first look at the numbers, it's clear that the Serbian students articulated more and more positive items (21 benefits and 11 challenges) compared to those given by students in Spain (11 benefits and 7 challenges).

FIGURE 2. Perceived benefits and challenges of ChatGPT use in learning.
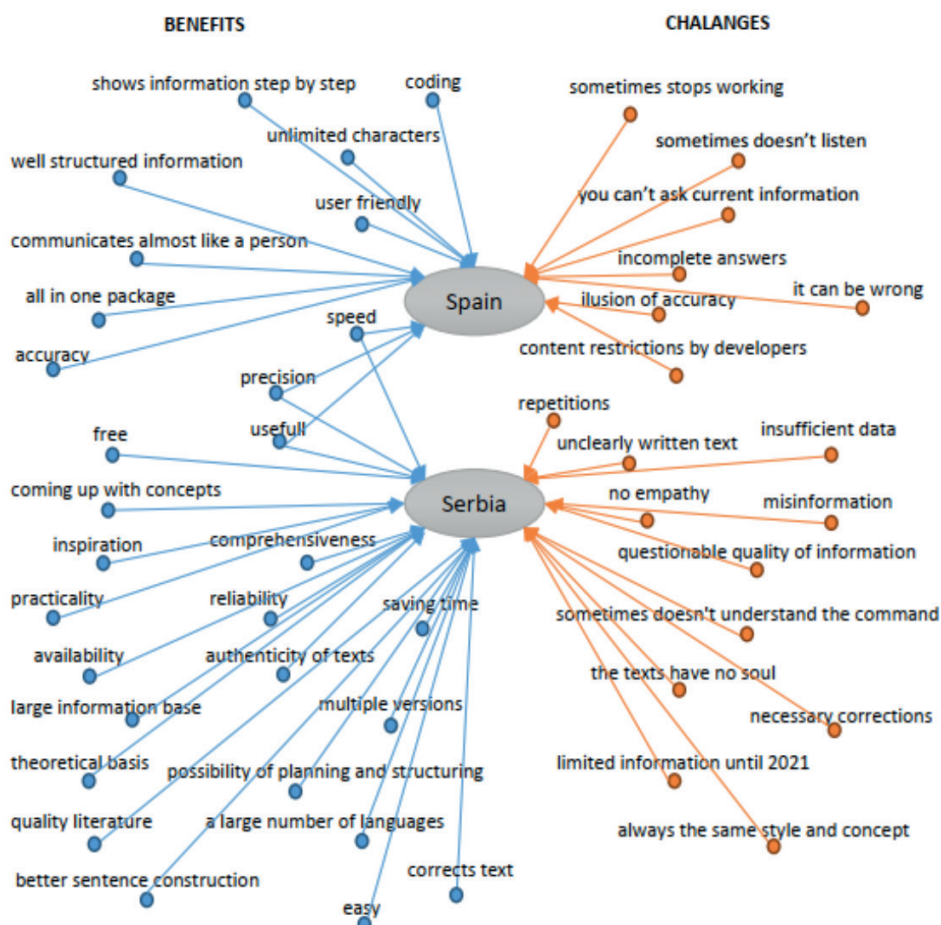
The results show that students from both countries have similar attitudes when it comes to the speed, accuracy and usefulness of ChatGPT. Namely, both groups singled out speed and precision, and then usefulness as important benefits in the application of ChatGPT in the learning process. They start from the position that this AI tool has the ability to quickly generate answers and provide accurate information, which is significant in the context of student learning. Positive answers include: "It does exactly what you are asking for and it's so fast" (E18) and points out that its most noticeable benefit is "absolutely incredible speed. I still can't believe that it is possible to get such precise and useful answers in just a few seconds" (S3).

Similarities in the answers are also present as regards to other characteristics of ChatGPT, such as for example the fact that it provides easy and well-structured information: "The information is well structured and usually in bullet points, which is easy to follow" (E8); "It was very useful for me when structuring the text, and it was very easy to find all the answers" (S7).

Differences in the perceived benefits among both groups appear on certain specific aspects. For example, students from Spain highlight that "it communicates almost like a person" (E19) and that "ChatGPT shows information step by step" (E17). Half of the Serbian students pointed out the fact that it is free as a benefit: "The biggest benefit is free. Most of the other tools I used have to be paid for" (S10). They also highlight its use as inspiration for ideas: "If I don't have an idea how to start work, I'll ask ChatGPT for inspiration and I'll start creating faster, and this way I'd suffer a creative block for days. It's good that it's free and available, I hope they don't start charging for it" (S4).

In line with the results of AlAfnan et al. (2023), a large number of students emphasise the benefit of easily obtaining a theoretical basis and the existence of a large database of information: "ChatGPT has benefits, first of all, speed; it enables quick data search which is complete, there is a lot of information available, it also uses a large number of languages and provides information in various languages" (S11); "It provides a theoretical basis in its proposals, which facilitates my further work and provides guidelines and sources for further research» (S19).

The differences among both groups can be explained by the existence of distinct priorities and preferences, which is arguably due to differences in their educational contexts, and in the goals and objectives of study programs. Students from Serbia cite benefits such as better sentence construction and authenticity of texts: "ChatGPT can construct my sentences better, can express my thoughts in a better way and generate an authentic text" (S13), among other answers related to writing style, which indicate that they have somewhat more specific needs related to the language and quality of the responses generated. Students in Serbia, here, are more focused on generating

textual content and stylistic language corrections. In contrast, those in Spain place a higher value on the technical aspects of working with this tool, particularly as regards to coding: "For example, if you search for code, it generates the exact code you are looking for. ChatGPT explains it to you, so you can understand it better" (E16).

As regards to the challenges, it is noticeable that both groups of students identified a smaller number of them compared to the number of identified benefits. They noticed problems related to reliability and operation: "Sometimes stops working" (E18); "Sometimes doesn't listen" (E18); "Sometimes doesn't understand the command" (S1). A significant challenge perceived by both groups relates to its limitations in providing current or up-to-date information. Namely, the students rightly stated that the currently available version of ChatGPT 3.5 has a limited information base, and it cannot provide data beyond what it was trained on: "It doesn't have internet connection, so you can't ask for current information' (E16); "The big drawback of ChatGPT is that it is limited to information until September 2021 (if I'm not mistaken), and there is no information after that date" (S12).

A very significant challenge of ChatGPT in the learning process, especially in the part of writing texts, was also noticed by both groups of students, and it refers to getting wrong or incomplete answers that at first glance seem correct: "The illusion of it having high accuracy of the information provided can be harmful since users may forget that it's a new tool that keeps improving every day. For example, when I asked about the battle of Kosovo that took place in 1389, it wrote to me about the wars of the nineties" (S8). This deficiency is especially pronounced when it comes to searching references for writing texts, where, according to the students, it happened that ChatGPT offered them a complete reference list that included non-existing texts: "Several times it gave me a list of references that didn't exist at all, especially when I was looking for more recent references from the social sciences" (S7). Accordingly, students feel that some texts generated by this tool are insufficiently credible, imprecise and can cause confusion and distraction: "When it comes to social and historical facts, that's when the greatest mistrust occurs" (S1).

As said above when describing ChatGPT's challenges, the students in Spain focus mostly on technical aspects and the time limitation and accuracy of the information. In contrast, students from Serbia, in addition to these technical aspects, cite other challenges that are more specific to their educational context, i.e., tasks related to studying and learning that they can implement with the help of this AI tool. These specific concerns are primarily related to the language and style of the generated content, because apparently students from Serbia receive such assignments to a greater extent, and consequently use

ChatGPT the most for writing different texts.

In this context, students from Serbia mention limited creativity and lack of imagination as one of the perceived challenges and notice that ChatGPT often formulates the generated text in almost the same way: "I think its creativity is limited and maybe not imaginative enough" (S4); "A problem can arise if two people ask to paraphrase the same passage, because they will give a very similar answer and there is no way to check this. The texts are very similar to each other, and the style and concept of writing is always the same" (S16). In addition, they point out that the texts generated by ChatGPT can be unclear or repetitive and also that the lack of an emotional component in the generated text can be noticed: "The texts it writes have no soul. It is only good when he needs to list the facts" (S8); "It is quite dry, lacks empathy in the texts he has put out so far" (S10).

In addition, students from Serbia state that there are challenges related to the limitation of content when answers are requested in languages other than English. "I've been using it for two months now. For data in English, it is fine, but for data in Serbian, it has no credibility' (S15). In relation to this, they point to the thematic limitation of the content, which can cause problems when it comes to certain topics or types of information that it generates: "Inaccurate data for unsolved murder issues, socio-social topics, school

violence topics, etc." (S15); "Generally known information in the world is correct, but when it comes to local information that, for instance, regarding Serbia, Niš, the information was incorrect. For example, he did not give the correct answer about FC Radnički, and the club was famous in Europe in history" (S16).

4.3. Potential for personalised learning

Figure 3 shows the results of the survey of the opinions of students from Spain and Serbia regarding the potential of ChatGPT for personalised learning. By reviewing the data, it can be concluded that there are no major differences in students' opinions. Almost all students from both countries recognize the potential of ChatGPT for personalised learning. The only exception is one student from Spain who thinks that this function of ChatGPT is not an option: "I don't think it can be possible" (E4). Other responses can be divided into four categories: individualised approach, learning plan, detailed explanations and suggestions.

As regards to individualised approach, the students recognize the potential of ChatGPT so that students can learn at their own pace: "Students can find topics that interest them and get at least basic knowledge in a certain area for free; they can learn at their own pace and when they want and as much as they want" (S10). Students from Spain also highlight the ability to customise ChatGPT and generate content: "[It] fits your pacing and availability" E(20).

FIGURE 3. Opportunities for personalised learning with ChatGPT.
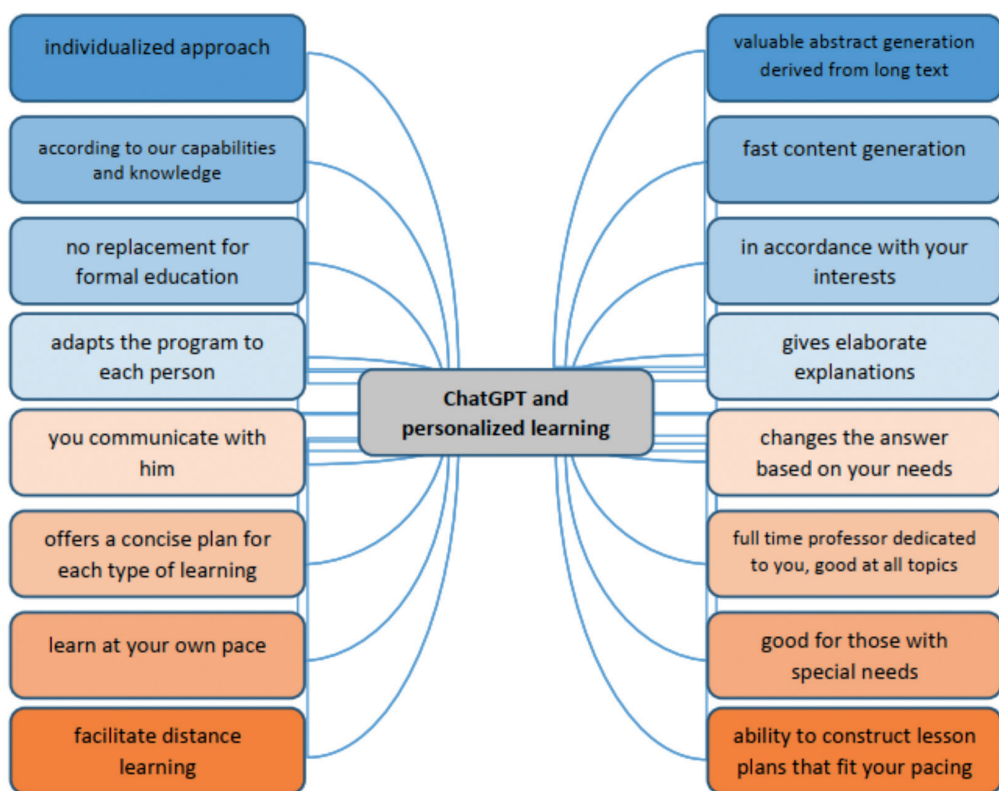
In addition, they state that ChatGPT matches the answers to cover a specific task: "For example, if you ask it for a code, you can change it based on your needs" (E16) and thus "facilitates distance learning through some exercises, examples, tests" (S13). They also recognize the usefulness of this AI tool for planning learning adapted to individual needs: "The great advantage is that it can, with good instructions, adapt the program to each person as it suits that person, and for any topic" (S7). Furthermore, it "offers a concise plan for each type of learning" (S9) and it is also generally good for any type of planning: "A very good function of this tool, because it is possible to create a plan (learning, diet, training...) which suits only you" (S12).

Students from both countries stated that ChatGPT provides detailed interpretations and answers to students' questions and requests related to the material through a personalised approach, "like having a professor dedicated to your questions full time and being good at all topics" (S15). According to them, this tool favours more effective learning because it "enables students to learn through interaction" (S11), to be more effective in learning because they can use "additional sources and sketches

of the material it creates" (S14), while receiving "elaborate explanations" (E8). Only one of the students stated that he saw "the potential of ChatGPT for the learning of persons with special needs" (E19).

However, along with this rather positive take on the topic, many students from both countries did show caution when it comes to using ChatGPT for personalised learning: "I believe that ChatGPT helps in the search for knowledge, but it is not adequate to completely replace the teacher and formal education or transfer experiences from person to person" (S6). In relation to improving the tool for this specific purpose, the students suggested it should "provide you videos or tutorials. That is an easier way to learn than just reading it" (E18).

5. Discussion and conclusions

The results presented above allow us, in line with our research goal, to paint a picture of the reception and intentions of future use of ChatGTP, and by extension similar AI tools, by university students. It also shows some interesting differences between the two groups that responded to the structured interview.

In summary, Spanish students demonstrated a more cautious approach to adopting ChatGPT for learning, with only 65% expressing intent to use it, unlike Serbian students, where 85% showed willingness. This discrepancy aligns with Boyon's (2022) research, suggesting that individuals from developing countries

tend to trust and adopt AI technologies more readily than those from developed nations.

Both groups of students agreed on using ChatGPT for organizing and creating presentations, as well as for writing extensive texts, paralleling Haensch et al.'s (2023) findings that students commonly employ ChatGPT for essays and coding tasks on social networks. Additionally, Biswas (2023a) noted ChatGPT's potential for proofreading tasks like grammar and style checks. However, there were important differences in opinions regarding proofreading between the two groups, with Serbian students displaying greater interest than their Spanish counterparts. This variation may stem from the more prevalent use and integration of Catalan and Spanish language tools in standard text applications, unlike Serbian, as suggested by student E15's response.

There was consensus on ChatGPT's advantages in the learning process, particularly its speed, accuracy, and overall usefulness, aligning with Keles and Aydin's (2021) findings that university students value AI's usefulness, ease of use, and its capacity for innovation. However, the challenges identified varied between groups. Spanish students primarily cited technical issues like time constraints and information accuracy, while Serbian students highlighted problems related to text generation, language, and writing style. Additionally, the latter group noted limitations in creativity and imagination, reflecting concerns corroborated

by Uludag (2023) regarding ChatGPT's creative capacities. These differences in perceived challenges might be attributed to varying academic needs and tasks, or even broader educational program content.

Regardless of the educational context they come from, students from both countries recognized the potential of ChatGPT for personalised learning. This is in fact where the greatest level of similarity between both groups is found, as they all show to perceive that ChatGPT has high potential to provide answers in accordance with individual needs, interests and individual knowledge level. They do believe, however, that it is not able to replace the teacher, a result that was also reached by Ausat et al. (2023).

In general terms, it can be concluded that students from both countries agree that ChatGPT can be a useful tool to support learning, although in their answers there is doubt and caution in relation to its use. As said above, the differences in opinions can arguably be explained by a different socioeconomic context, their different educational contexts, and also by the availability of other useful digital tools, their individual preferences and the level of information about AI technology.

Despite the limitations of our small sample size and the recent emergence of ChatGPT, our research offers valuable initial insights into the use of this AI tool across two countries. These findings aim to contribute to the ongoing global discussion on AI technologies in education, emphasizing the need for increased dialogue and collaboration among educational institutions, researchers, and experts from diverse fields and regions to enhance the effectiveness and meaningful integration of AI in learning environments.

Future work includes expanding our sample into a wider number of respondents and the presence of more countries. It also includes iterating the design of the written structured interview in order to obtain responses that help towards a more in depth understanding of the perceptions and intentions of use of AI tools in general, and of how the differences among different cultural and educational contexts can influence the acceptance of artificial intelligence technologies.

Given the projected significance of AI in education over the next two decades (Zawacki-Richter et al., 2019), research into student perceptions and the broader application of AI tools is key. Such studies do not only enrich our understanding but are very important for integrating AI effectively into higher education curricula (Neumann et al., 2023). While the full impact of tools like ChatGPT on educational practices remains uncertain, it is essential to acknowledge the significant role educators and developers play in responsibly leveraging these technologies to enhance learning outcomes.

Authors' contributions

Dragana Pavlovic: Conceptualisation; Data curation; Formal analysis; Methodology; Writing (original draft); Writing (revision and editing); Visualisation.

Joan Soler-Adillon: Conceptualisation; Data curation; Funding acquisition; Writing (revision and editing); Validation.

Zorica Stanisavljevic-Petrovic: Conceptualisation; Data curation; Formal analysis; Funding acquisition; Methodology; Supervision.

Funding

References

Addington, S. (2023, May 9). ChatGPT: Cyber security threats and counter-measures. *SSRN*. http://dx.doi.org/10.2139/ssrn.4425678

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15* (3), ep429. https://doi.org/10.30935/cedtech/13152

AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, *3* (2), 60-68. https://doi.org/10.37965/jait.2023.0184

Alshater, M. M. (2023, January 4). Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT. *SSRN*. http://dx.doi.org/10.2139/ssrn.4312358

Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI*. DigitalCommons@URI, University of Rhode Island. https://digitalcommons.uri.edu/cba_facpubs/548

Ausat, A., Massang, B., Efendi, M., Nofirman, N., & Riady, Y. (2023). Can ChatGPT replace the role of the teacher in the classroom: A fundamental analysis. *Journal on Education*, *5* (4), 16100-16106. https://jonedu.org/index.php/joe/article/view/2745

Avila-Chauvet, L., & Mejía, D. (2023, March 4). Can professors and students detect ChatGPT essays? *SSRN*. http://dx.doi.org/10.2139/ssrn.4373643

Baidoo-Anu, D., & Ansah, O. L. (2023, April 13). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN*. http://dx.doi.org/10.2139/ssrn.4337484

Baskara, F. R. (2023). The promises and pitfalls of using ChatGPT for self-determined learning in higher education: An argumentative review. *Prosiding Seminar Nasional Fakultas Tarbiyah dan Ilmu Keguruan IAIM Sinjai*, *2*, 95-101. https://doi.org/10.47435/sentikjar.v2i0.1825

Biswas, S. (2023a, March 31). The function of ChatGPT in social media: According to ChatGPT. *SSRN*. http://dx.doi.org/10.2139/ssrn.4405389

Biswas, S. (2023b, June 4). Role of ChatGPT in education. *SSRN*. https://ssrn.com/abstract=4369981

Borsos, É., Banos-González, I., Boric, E., Lyngved, R., & Fekete, A. B. (2022). Trainee teachers' perceptions of outdoor education. *Environmental Education Research*, *28* (10), 1490-1509. https://doi.org/10.1080/13504622.2022.2031901

Boyon, N. (2022). Opinions about AI vary depending on countries' level of economic development. *Ipsos*. https://www.ipsos.com/en/global-opinions-about-ai-january-2022

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, *18* (1), 198-204. https://doi.org/10.5281/zenodo.7716416

Brkanlić, S., Sánchez-García, J., Esteve, E. B., Brkić, I., Ćirić, M., Tatarski, J., Gardašević, J., & Petrović, M. (2020). Marketing mix instruments as factors of improvement of students' satisfaction in higher education institutions in Republic of Serbia and Spain. *Sustainability, 12* (18), 7802. https://doi.org/10.3390/su12187802

Corbí, M., Tombolato, M., Bueno-Sánchez, L., Hermans, K., Valenti, A., Garcés, J., Straniero, A. M., Brojčin, B. B., Mesquita, C., Bonifácio, E., Martini, B., Rodríguez, S., Milants, M., Glumbić P., & Glumbić, N. (2021). Intellectual disability in higher education: Self-perceived training needs of university teachers. *Specijalna edukacija i rehabilitacija*, *20* (1), 1-15. https://doi.org/10.5937/specedreh20-31002

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61* (2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Dai, W., Lin, J., Jin, F., Li, T., Tsai, Y., Gasevic, D., & Chen, G. (2023, July 5). Can large language models provide feedback to students? A case study on ChatGPT. *EdArXiv*. https://doi.org/10.35542/osf.io/hcgzj

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40* (2), 615-622. https://doi.org/10.5114/biolsport.2023.125623

Despotovic, D., Cvetanovic, S., Nedic, V., Despotovic, M. (2019). Social aspects of sustainable competitiveness in the selected European countries in the period 2012-2015. *Social Indicators Research*, *141*, 841-860. https://doi.org/10.1007/s11205-018-1840-4

Dhiman, B. (2022, 23 de octubre). Artificial intelligence and voice assistant in media studies: A critical review. *SSRN*. http://dx.doi.org/10.2139/ssrn.4250795

Dolenec, D., Balković, A., Kralj, K., Širinić, D., Romanos, E., & Fernandes, T. (2020). Protest event dataset for Croatia, Portugal, Serbia and Spain: Focus on strike data. *Politička misao: časopis za politologiju*, *57* (4), 155-168. https://doi.org/10.20901/pm.57.4.07

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., […] y Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Eze, S. C., Chinedu-Eze, V., & Bello, A. O. (2018). The utilisation of e-learning facilities in the educational delivery system of Nigeria: a study of M-University. *International Journal of Educational Technology in Higher Education*, *15* (34), 1-20. https://doi.org/10.1186/s41239-018-0116-z

Ezquerra, A., Agen, F., Rodríguez-Arteche, I., Ezquerra-Romano, I. (2022). Integrating artificial intelligence into research on emotions and behaviors in science education. *Eurasia Journal of Mathematics, Science and Technology Education*, *18* (4), em2099. https://doi.org/10.29333/ejmste/11927

Ezquerra, A., Agen, F., Toma, R. B., Ezquerra-Romano, I. (2023). Using facial emotion recognition to research emotional phases in an inquiry-based science activity. *Research in Science & Technological Education*, 1-24. https://doi.org/10.1080/02635143.2023.2232995

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *61* (3), 460-474. https://doi.org/10.1080/14703297.2023.2195846

Fedorov, A., & Levitskaya, A. (2015). The framework of Media Education and Media Criticism in the Contemporary World: The opinion of International Experts. (Situación de la educación en medios y la competencia crítica en el mundo actual: opinión de expertos internacionales). *Comunicar*, *23* (45), 107-125. http://dx.doi.org/10.3916/C45-2015-11

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. J*ournal of Applied Learning and Teaching*, *6* (1), 1-7. https://doi.org/10.37074/jalt.2023.6.1.22

Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching*, *12* (1), 44-58. https://doi.org/10.15294/elt.v12i1.64069

Frye, B. L. (2022). Should using an AI text generator to produce academic writing be plagiarism? *Fordham Intellectual Property, Media & Entertainment Law Journal, Forthcoming*, *33* (4), 5. https://ir.lawnet.fordham.edu/iplj/vol33/iss4/5

García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society*, *24*, e31279. https://doi.org/10.14201/eks.31279

Haensch, A. C., Ball, S., Herklotz, M., & Kreuter, F. (2023, March 9). Seeing ChatGPT through students' eyes: An analysis of TikTok data. *arXiv:2303.05349v1*. https://doi.org/10.48550/arXiv.2303.05349

Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, *15*(2), ep421. https://doi.org/10.30935/cedtech/13036

Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2* (4), 100089. https://doi.org/10.1016/j.tbench.2023.100089

Hariri, W. (2023, April 17). Unlocking the potential of ChatGPT: A comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing. *arXiv:2304.02017v10*. https://doi.org/10.48550/arXiv.2304.02017

Health, T. L. D. (2023). ChatGPT: Friend or foe? *The Lancet: Digital health*, *5* (3), e102. https://doi.org/10.1016/S2589-7500(23)00023-7

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. In C. Stückelberger, & P. Duggal (Eds.), *Data ethics: Building trust. How digital technologies can serve humanity [Ética de los datos: construir la confianza. Cómo las tecnologías digitales pueden servir a la humanidad]* (pp. 621-653). Globethics Publications. https://doi.org/10.58863/20.500.12424/4276068

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5* (1), 37-45. https://jeti.thewsu.org/index.php/cieti/article/view/103/64

Iskender, A. (2023). Holy or unholy? Interview with Open AI's ChatGPT. *European Journal of Tourism Research*, *34*, 3414-3414. https://doi.org/10.54055/ejtr.v34i.3169

Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023, November 2). Is ChatGPT a good translator? A preliminary study. *arXiv:2301.08745v4*. https://doi.org/10.48550/arXiv.2301.08745

Kalla, D., & Smith, N. (2023). Study and analysis of ChatGPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, *8* (3). https://doi.org/10.5281/zenodo.10250455

Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing*, *72* (12), 2954-2965. https://doi.org/10.1111/jan.13031

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Keles, P. U., & Aydin, S. (2021). University students' perceptions about artificial intelligence. *Shanlax International Journal of Education*, *9*, 212-220. https://doi.org/10.34293/education.v9iS1-May.4014

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My teacher is a machine: Understanding students' perceptions of AI teaching assistants in online education. *International Journal of Human-Computer Interaction*, *36* (20), 1902-1911. https://doi.org/10.1080/10447318.2020.1801227

Kitamura, F.C. (2023). ChatGPT is shaping the future of medical writing but still requires human judgment [ChatGPT está dando forma al futuro de la redacción médica, pero aún requiere criterio humano]. *Radiology*, *307* (2), e230171. https://doi.org/10.1148/radiol.230171

Klutka, J., Ackerly, N., & Magda, A. J. (2020). *Artificial intelligence in higher education. Current uses and future applications*. Learning House.

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M. D., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13* (18), 10424. https://doi.org/10.3390/su131810424

Lee, H. (2023, March 14). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education*. https://doi.org/10.1002/ase.2270

Lee, Y (2003). The technology acceptance model: Past, present and future. *Communications of the Association for Information Systems (CAIS)*, *12* (50), 752-780. https://doi.org/10.17705/1CAIS.01250

Mathers, N., Fox, N., & Hunn, A. (2002). *Trent Focus for research and development in primary health care: Using interviews in a research project*. Trent Focus Group.

Megahed, F. M., Chen, Y. J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023, February 17). How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study. *arXiv:2302.10916v1*. https://doi.org/10.48550/arXiv.2302.10916

Mollick, E. R., & Mollick, L. (2022, December 23). New modes of learning enabled by AI chatbots: Three methods and assignments. *SSRN*. http://dx.doi.org/10.2139/ssrn.4300783

Neumann, M., Rauschenberger, M., & Schön, E. M. (2023). "We need to talk about ChatGPT": The future of AI and higher education. In *2023 IEEE/ACM 5th International Workshop on Software Engineering Education for the Next Generation (SEENG)* (pp. 29-32). IEEE.

Opara, E., Adalikwu, T., & Aduke, T. C. (2023). ChatGPT for teaching, learning and research: Prospects and challenges. *Global Academic Journal of Humanities and Social Sciences*, 5 (2), 33-40. https://doi.org/10.36348/gajhss.2023.v05i02.001

OpenAI. (2023). *ChatGPT* (March 14th release) [Large Language Model]. https://chat.openai.com/chat

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator, 78* (1), 84-93. https://doi.org/10.1177/10776958221149577

Podstawski, R., Żurawik, M., Borysławski, K., Omelan, A. A., & Tsos, A. (2022). Working conditions of physical education teachers in European higher education institutions. *Physical Activity Review, 10* (2), 111-126. https://www.physactiv.eu/wp-content/uploads/2022/04/2022_10226.pdf

Qadir, J. (2023). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE. https://doi.org/10.36227/techrxiv.21789434.v1

Raman, R., Mandal, S., Das, P., Kaur, T., Sanjanasri, J. P., & Nedungadi, P. (2023). *University students as early adopters of ChatGPT: Innovation diffusion study* [Unpublished]. https://doi.org/10.21203/rs.3.rs-2734142/v1

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, *6* (1), 1-16. https://doi.org/10.37074/jalt.2023.6.1.29

Rodrigues, M. A., & dos Santos, P. R. (2023, May 1). Enhancing STEM learning with ChatGPT and Bing Chat as objects-to-think-with: A case study. *arXiv:2305.02202v1*. https://doi.org/10.48550/arXiv.2305.02202

Romero-Rodríguez, J., Ramírez-Montoya, M., Buenestado-Fernández, M., & Lara-Lara, F. (2023). Use of ChatGPT at university as a tool for complex thinking: Students' perceived usefulness. *Journal of New Approaches in Educational Research*, *12* (2), 323-339. https://doi.org/10.7821/naer.2023.7.1458

Rose, R. (2023). *ChatGPT in higher education. Artificial intelligence and its pedagogical value*. Pressbooks.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6* (1). https://doi.org/10.37074/jalt.2023.6.1.9

Russell, S., & Norvig, P. (2010). *Artificial intelligence: A modern approach (3rd ed.)*. Prentice Hall

Sabzalieva, E., & Valentini, A. (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. IESALC. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

Schiff, D. (2021). Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI & Society, 36*, 331-348. https://doi.org/10.1007/s00146-020-01033-8

Sebastian, G. (2023, February 24). Do ChatGPT and other AI chatbots pose a cybersecurity risk? An exploratory study. *SSRN*. http://dx.doi.org/10.2139/ssrn.4363843

Silva, L. (2007). Post-positivist review of technology acceptance model. *Journal of the Association for Information Systems*, *8* (4), 255-266. https://doi.org/10.17705/1jais.00121

Sison, A. J. G., Daza, M. T., Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023, April 6). ChatGPT: More than a weapon of mass deception. Ethical challenges and responses from the human-centered artificial intelligence (HCAI) perspective. *arXiv:2304.11215v1*. https://doi.org/10.48550/arXiv.2304.11215

Sok, S., & Heng, K. (2023, 22 de mayo). ChatGPT for education and research: A review of benefits and risks. *SSRN*. http://dx.doi.org/10.2139/ssrn.4378735

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching, 6* (1), 31-40. https://doi.org/10.37074/jalt.2023.6.1.17

Susnjak, T. (2022). ChatGPT: The end of online exam integrity? *arXiv:2212.09292v1*. https://doi.org/10.48550/arXiv.2212.09292

Tian, H., Lu, W., Li, T.O., Tang, X., Cheung, S.-C., Klein, J., & Bissyandé, T. F. (2023, August 31). Is ChatGPT the ultimate programming assistant - How far is it? *arXiv:2304.11938v2*. https://arxiv.org/abs/2304.11938

Tiunova, A., & Muñoz, F. (2023, June 28). ChatGPT: Using AI in social studies academic research. *SSRN*. http://dx.doi.org/10.2139/ssrn.4451612

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments 10*, 15. https://doi.org/10.1186/s40561-023-00237-x

Uludag, K. (2023, March 22). Testing creativity of ChatGPT in psychology: Interview with ChatGPT. *SSRN*. http://dx.doi.org/10.2139/ssrn.4390872

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19* (7), em2286. https://doi.org/10.29333/ejmste/13272

Woo, J. H., & Choi, H. (2021, October 21). Systematic review for AI-based language learning tools. *arXiv:2111.04455v1*. https://doi.org/10.48550/arXiv.2111.04455

Woolf, B. P., Lane, H. C., Chaudhri, V. K., & Kolodner, J. L. (2013). AI grand challenges for education. *AI magazine, 34* (4), 66-84. https://doi.org/10.1609/aimag.v34i4.2490

Yu, H. (2023). Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, *14*, 2156. https://doi.org/10.3389/fpsyg.2023.1181712

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, (16), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2023, June 4). ChatGPT user experience: Implications for education. *SSRN*. http://dx.doi.org/10.2139/ssrn.4312418

Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023, May 29). Exploring AI ethics of ChatGPT: A diagnostic analysis. *arXiv:2301.12867v4*. https://arxiv.org/abs/2301.12867

Authors' biographies

Dragana Pavlovic. She is an associate professor at the Department of Communication and Journalism at the Faculty of Philosophy, University of Nis, Serbia. She holds two PhDs: one in Technical Sciences (Telecommunications), and another in Social Sciences and Humanities (Communication and Education). She has published two books and more than 60 research papers. Her research primarily focuses on new media and education.

iD http://orcid.org/0000-0002-5115-8698

Joan Soler-Adillon. He is an associate professor at the Faculty of Computer Science, Multimedia and Telecommunications at the Universitat Oberta de Catalunya (UOC), in Barcelona (Spain), where he is the director of the research group DARTS. He has previously held academic positions at Royal Holloway (University of London)

and at Universitat Pompeu Fabra. His research and practice revolve around digital interactive media and its use in arts, digital creativity, and education.

https://orcid.org/0000-0002-7959-4192

Zorica Stanisavljevic-Petrovic. She is full professor at the Department of Pedagogy, Faculty of Philosophy, at the University of Nis, Serbia. She received her PhD in Pedagogy from the Faculty of Philosophy, Belgrade, Serbia. She has published 5 books and more than 80 papers in national and international journals and proceedings in the area of school theories, institutional upbringing and education, school problems, new media and education.

http://orcid.org/0000-0001-9212-4286

Teachers' attitudes towards the use of virtual simulations: Design and validation of a questionnaire

Actitud del profesorado hacia el uso de simulaciones virtuales: diseño y validación de un cuestionario

Alicia PALACIOS, PhD. Associate Professor. Universidad Internacional de La Rioja (UNIR), Spain (*alicia.palacios@unir.net*).

Rosa GÓMEZ, PhD. Associate Professor. Universidad Internacional de La Rioja (UNIR), Spain (*rosa.gomez@unir.net*).

Álvaro BARRERAS, PhD. Associate Professor. Universidad Internacional de La Rioja (UNIR), Spain (*alvaro.barreras@unir.net*).

Daniel MORENO-MEDIAVILLA, PhD. Associate Professor. Universidad Internacional de La Rioja (UNIR), Spain (*daniel.moreno@unir.net*).

Abstract:

Virtual simulations in science classes are not used as often as might be expected if we consider the demonstrated improvement in students' conceptual learning, the development of their skills, and their acquisition of positive emotions. This makes it necessary to identify teachers' attitudes and perceptions to the use of these tools. The aim of this work is to construct and validate an instrument for measuring the attitudes of secondary education teachers towards the use of virtual simulations in STEM fields. Based on an in-depth theoretical review, we developed an initial questionnaire that was subjected to a process of expert validation and a pilot study. A questionnaire comprising 27 items was obtained, which was applied to 783 secondary school teachers in Spain. After carrying out confirmatory factor analysis, a scale comprising five factors was obtained. The psychometric analyses displayed satisfactory fit indices that prove the discriminant and convergent validity of the model. The result is a useful instrument for determining the principal factors that discourage teachers from habitually using simulations. This enables the design of training proposals that take teachers' prior attitudes into account.

rep

Keywords: virtual simulations, teacher attitude, questionnaire, secondary teaching.

Resumen:

Las simulaciones virtuales en el aula de ciencias no se utilizan de manera tan habitual como cabría esperar si se tiene en cuenta la mejora demostrada en el aprendizaje conceptual del estudiante, en el desarrollo de sus habilidades y en la adquisición de emociones positivas. Esto hace necesario conocer cuáles son las actitudes y percepciones del profesorado en el uso de estas herramientas. El objetivo de este trabajo es construir y validar un instrumento de medida de las actitudes del profesorado de educación secundaria hacia el uso de simulaciones virtuales de áreas STEM. A partir de una profunda revisión teórica, se desarrolló un cuestionario inicial, que fue sometido a un proceso de validación de expertos, y un estudio piloto. Se obtuvo un cuestionario formado por 27 ítems, el cual fue aplicado a 783 profesores de educación secundaria de España. Tras el análisis factorial confirmatorio desarrollado, se obtuvo una escala compuesta por cinco factores. Los análisis psicométricos mostraron índices de ajuste satisfactorios que prueban la validez discriminante y convergente del modelo. El resultado es un instrumento útil para determinar los factores principales que alejan al profesorado del uso habitual de las simulaciones. Esto posibilita el diseño de propuestas de formación que tengan en cuenta las actitudes previas del docente.

Palabras clave: simulaciones virtuales, actitud del docente, cuestionario, enseñanza secundaria.

1. Introduction

In current science teaching focuses, the learning of processes and understanding of content are becoming ever more important. Given this need, new technological challenges that favour a holistic learning of science are emerging (Oliverira et al., 2019; Osborne, 2014). In this sense, virtual simulations (VS) offer strategies based on the scientific method (De Jong & Van Joolingen, 1998; Fan & Geelan, 2013; Rutten et al., 2012), which favour reasoning and inquiry competences (Stieff, 2019; Trujillo et al., 2023; Wen et al., 2020), graphical competences (Plass et al., 2012), higher-order thinking skills (Amin & Ikhsan, 2021), and scientific literacy (Chen et al., 2014; Lynch & Ghergulescu, 2017). Furthermore, they have the capacity to support model-based learning and its effect on visualisation (Lee et al., 2021) of most of the content associated with STEM (science, technology, engineering, and mathematics) fields (D'Angelo et al., 2014). They also make it possible to represent real phenomena: students observe and manipulate variables and phenomena, and visualise the changes that occur (Chan et al., 2021; De Jong & Van Joolingen, 1998). Finally, the use of virtual simulations improves students' satisfaction and engagement (Durán et al., 2007; Wu & Huang, 2007), and so they

develop positive attitudes towards the sciences (Zacharia, 2003).

All of these characteristics make the introduction of virtual simulations as a didactic resource very beneficial (D'Angelo et al., 2014; Waight et al., 2014). However, they are not habitually used in classrooms (Chan et al., 2021; Lee et al., 2021). To resolve this situation, a need to work from two fundamental perspectives emerges: teachers' knowledge of the use of virtual simulations (Gómez et al., 2022; Moreno-Mediavilla et al., 2023) and their attitude towards using them (Lee et al., 2021). The aim of the present work relates to this latter perspective.

Attitude is described as an individual's favourable or unfavourable outlook on an object, a person, or an event (Albirini, 2006). This is a multidimensional construct that encompasses three dimensions (affective, behavioural, and cognitive) although their inclusion and interpretation vary according to the authors. For example, Teo et al. (2016) include the affective dimension, understood as teachers' enjoyment of, pleasure from, and preference for the use of technology. Shapka and Ferrari (2003) also add the factor of anxiety when using technology. Not all authors regard the behavioural dimension as relevant. However, Teo (2008) does consider it to be important, as it makes it possible to analyse how proactive teachers are towards different technological resources, their regular use, and training. The cognitive dimension is associated with aspects such as perceived time savings or the improvements that technology provides in

the educational process (Albirini, 2006). Other cognitive components are also included, such as teachers' perception of the usefulness of resources or their skills in using them (Teo, 2008), or the beliefs, valuation, and social function that they give to the use of technology (Cai et al., 2017).

Based on these definitions, many authors have analysed the principal variables that affect the use of technology in education, drawing on the TAM theoretical model (technology acceptance model) (Davis et al., 1989). According to this model, the perceived usefulness and ease of use of technology significantly influence attitudes towards using it, and this is related to the intention to use the technology and its actual use (Albirini, 2006; Tate et al., 2015; Teo et al., 2016). In addition to these two principal factors, other possible variables have been described: the conditions that facilitate the use of technology (Lai et al., 2012; Teo, 2012), technological complexity (Teo, 2012), teacher self-efficacy (Wong et al., 2012), gender (Cai et al., 2017), age (Nunes et al., 2020), years of experience of use (Gargallo et al., 2007), and field of study (Teo, 2008).

The extensive bibliography on attitudes towards the use of technology contrasts with the few works on attitudes towards the specific use of virtual simulations. Zacharia (2003) noted that teachers' attitudes towards the use of technology improved after using it in class. For their part, Lehtinen et al. (2016) showed that knowledge of technology is correlated with the perceived usefulness of simulations and with attitudes towards their use in

class. Recently, Lee et al. (2021) analysed teachers' attitudes towards the use of simulations through interviews. They found that the principal variables were perceived usefulness, ease of use, appropriateness of the simulations, the possibility of engaging students in learning, and the teacher's professional growth needs. With regards to the use of virtual simulations by teachers, Gómez et al. (2022) and Moreno-Mediavilla et al. (2023) designed and administered a questionnaire to analyse the perception of teachers' perception of their digital competence in the actual use of virtual simulations. The questionnaire presented in this work provides information of great interest for addressing teachers' specific attitudes towards the use of virtual simulations, as there is no specific instrument for evaluating them. The attitudes and beliefs of the teacher on the use of these tools are decisive for improving their willingness to use them (Khan, 2011; Padilla, 2018). As well as identifying teachers' attitudes towards the use of virtual simulations, this questionnaire establishes whether this attitude directly influences the use of such tools and compares it with teachers' own perceptions of their digital competences. This question has been highlighted as one of the principal needs in teacher training, as noted by De Pro Bueno et al. (2022) and Pozuelo et al. (2023).

Given the importance of knowing teachers' attitudes towards the implementation of virtual simulations in the classroom, this work aims to design and validate a questionnaire on secondary teachers' attitudes towards the use of virtual simulations in STEM.

2. Method

2.1. Research design

This study is quantitative, non-experimental, and descriptive and uses a transversal design.

2.2. Sample

The sample was obtained using non-probability convenience sampling. It consists of 783 secondary education teachers from Physics, Chemistry, Biology, Geology, Technology, and Mathematics from all over Spain, who work in public (80.72 %), state assisted (16.22 %), and private (3.07 %) schools in urban (73.56 %) and rural (26.44 %) areas. Of the participants, 61.94 % are women and 37.68 % men. In relation to age, 8.43 % of the teachers are aged under 31; 20.05 %, between 31 and 40; 37.42 %, between 41 and 50; 31.42 %, between 51 and 60, and 2.68 % of them are more than 60 years old. With regards to the distribution of years of teaching experience, 22.09 % have less than 5 years of experience, 17.24 % have between 5 and 10 years, 24.39 % have between 11 and 20 years, and 36.27 % have more than 20 years' teaching experience.

2.3. Process of preparation of the questionnaire

The process of preparing this instrument takes as its starting point, the TAM model of Davis et al. (1989), who suggest that the usefulness and ease of use of technologies directly influence people's attitudes and, with it, their final behaviour. Therefore, the questionnaire proposes an initial model with two major factors (usefulness and ease of use). In addition, other

variables that might influence the use of technologies and, in particular, virtual simulations were taken into account for the design of the items (Lee et al., 2021, Albirini, 2006, Teo et al., 2016, Tate et al., 2015). So, items related to the affective dimension of the attitude towards the use of virtual simulations were proposed, like those previously described by Teo et al. (2016). Likewise, items were defined that relate to perceived usefulness and are widely described by various authors (Albirini, 2006; Sahin et al., 2016; Teo, 2008; Teo et al., 2016). These items were proposed considering two additional aspects: their usefulness for improving understanding of content and their usefulness for acquiring scientific competences. Finally, in line with the need proposed by Lee et al. (2021) to consider in depth the aspects involved in teachers' perceptions of the ease of use of simulations, items that relate to their availability, complexity of use, and possible obstacles for students' learning were introduced.

Through this analysis, 46 relevant items were identified for the first version of the instrument. With the aim of determining the content validity of the questionnaire, a process of expert judgement was carried out by a panel of 4 teachers and researchers [an adequate number according to Grant and Davis (1997)] selected by convenience. These professionals had a variety of professional experience (from 10 to 46 years) and were experts of recognised standing in the didactics of experimental sciences, didactics of mathematics, and research methods in education. The judgement they provided included analysis of

each item in relation to three properties: clarity, coherence, and relevance, as well as the relevant comments for each case. In cases where two or more experts negatively evaluated the coherence or relevance properties, the item was eliminated; when clarity was criticised, the item's wording was reviewed.

A pilot trial was then carried out with a sample of 30 secondary education teachers from STEM fields in Spain, selected by convenience sampling; the objective of this was to obtain information about the relevance of the items. The teachers were from public and state-assisted centres in urban and rural locations. The questionnaire was distributed by email during the first two weeks of October 2021. A Cronbach's alpha analysis of the questionnaire as a whole was performed and the item-total correlation was analysed.

2.4. Instrument and data collection procedure

The CADUSV questionnaire (Teacher Attitudes to the Use of Virtual Simulations Questionnaire), drawn up using the Survey Monkey tool, includes several questions to characterise the sample (sex, age, years of teaching experience, speciality, and type of educational centre) and 27 items relating to the attitudes of STEM teachers towards the use of virtual simulations (link to the questionnaire). The teachers indicate their level of agreement or disagreement with each item on a Likert scale from 1 to 5. In addition, the objective of the study is set out in the instrument and the participation of teachers of Physics and Chemistry, Biology and

Geology, and Technology and Mathematics in secondary education is requested.

The questionnaire was distributed manually by email to public, state-assisted, and private educational centres in Spain. It was distributed between the second half of January and the end of February 2022 during which period participation in this study was active. Participants' anonymity was ensured at all moments.

2.5. Data analysis

The existence of previous studies directly related to the topic made it possible to start from a series of factors when constructing the questionnaire. Therefore, the construct validation was done through confirmatory factor analysis (CFA), for which the Mplus 8.1.5 data analysis program was used.

Using the SPSS 25 program, the samples were first checked for the presence of outliers; any found were eliminated from the total sample. The assumption of multivariate normality was tested using the Mardia coefficient (1970), which is determined based on the skew and kurtosis figures. As this assumption was not fulfilled, the polychoric matrix was used as the basis of the CFA. With the aim of fully rejecting the option of using the Pearson matrix as the basis of the analysis, tests were carried out with both types of matrix. In the case of the Pearson matrix, worse fit figures were observed. The estimate of the models was developed using the unweighted least squares mean and variance adjusted (ULSMV) estimator, given that the variables observed for the proposed model were ordinal.

The indices of fit of the two-dimensional model were calculated first and the modification indices and standardised residuals were obtained with the aim of improving the fit of the models. Absolute fit indices, which make it possible to evaluate whether the underlying theory fits the data extracted, were used as well as incremental fit indices, which compare the proposed model to a base model (McNeish et al., 2018).

The following absolute fit indices were obtained: the chi-squared value divided by the degrees of freedom for the model (CMIN/*df*), where values below 2 are very good and between 3 and 5 are acceptable (Hair et al., 2014); and the root mean square error of approximation (RMSEA), which seeks to test the fit between the proposed model and hypothetical populational data. In this case, values below .05 are considered very good, and between .05 and .08, acceptable (Byrne, 2009). The incremental fit indices provided are CFI (comparative fit index) and TLI (Tucker–Lewis index). In these, values greater than .9 are adequate and from .95 they are considered optimal (Xia & Yang, 2019).

With the aim of determining the discriminant validity, the two-factor model was analysed and was compared with the single-factor model, as well as with the four, five, and six factor models. The existence of second-order models was also considered (Table 1). These alternative models were defined on the basis of different possible combinations that were consistent with the theoretical assumption described above. For the comparison of the models, the indices of fit specified above were studied.

TABLE 1. Organisation of the different models regarding items and factors.

Item	2f model	4f model	5f model	6f model	5f model with second-order factors	6f model with second-order factors
1						
2						
3			Perceived usefulness (PU)	Perceived usefulness for conceptual learning (PUConc)	Perceived usefulness (PU)	Perceived usefulness for conceptual learning (PUConc)
4						
5						
6						
7	Interest in use (IU)	Interest in use (IU)			Factor 1: interest in use (IU)	Factor 1: interest in use (IU)
8				Perceived usefulness for competence learning (PUComp)		Perceived usefulness for competence learning (PUComp)
9						
10						
11						
12*			Beliefs and feelings about the use of VS (BFU)	Beliefs and feelings about the use of VS (BFU)	Beliefs and feelings about the use of VS (BFU)	Beliefs and feelings about the use of VS (BFU)
13						
14						

Item					
15					Complexity of use (CU)
16	Complexity of use (CU)	Complexity of use (CU)	Complexity of use (CU)		Complexity of use (CU)
17		Complexity of use (CU)	Complexity of use (CU)		Complexity of use (CU)
18			Complexity of use (CU)		Complexity of use (CU)
19	Perceived difficulty (PD)			Factor 2: perceived difficulty (PD)	Factor 2: perceived difficulty (PD)
20			Factor 2: perceived difficulty (PD)		
22	Obstacles for learning (OL)	Obstacles for learning (OL)	Obstacles for learning (OL)		Obstacles for learning (OL)
23			Obstacles for learning (OL)		Obstacles for learning (OL)
25	Availability (A)	Availability (A)	Availability (A)		
26	Availability (A)	Availability (A)	Availability (A)		Availability (A)
27	Availability (A)	Availability (A)	Availability (A)		Availability (A)

*Item 12, in the initial analyses, belongs to the *perceived difficulty* dimension, as explained in the results section.

Also, to determine the shared variance of the items that form a dimension the convergent validity was analysed by analysing the factor loadings, which must achieve values greater than .5. Within the convergent validity, composite reliability was also analysed (Green & Yang, 2015) as well as the average variance extracted (AVE). To be considered acceptable, these values must be above .7 and .37-.5, respectively (Hair et al., 2014; Moral, 2019).

3. Results

The initial model proposed for validation comprised 27 observed variables and two latent dimensions: interest in use (IU) and perceived difficulty (PD) (Table 1).

The analysis of outliers revealed the existence of six data sets that had to

eliminated, and so the final study sample comprised 777 teachers, 99% of the initial sample. After establishing that the assumption of multivariate normality of Mardia was not fulfilled (Table 2) and that the fit with the Pearson matrix did not generate satisfactory results, it was decided to use the polychoric correlation matrix for the CFA.

We initially started from the estimate of the model of two dimensions (IU and PD) to then improve it based on the R^2 values and the modification indices of each item. As Table 3 shows, almost all of the R^2 values were greater than .5, and so they were considered to be items that were adequate for explaining the model. However, it was decided that items 21 and 24, with R^2 values close to zero, did not have explanatory power and so they were eliminated from the questionnaire.

TABLE 2. Mardia analysis.

	Coefficient	Statistic	*df*	*p*
Skew	67.269	8711.274	3654	1.000
Adjusted skew for small samples	67.269	8.747.318	3654	1.000
Kurtosis	924.526	49.845		.000*

Note: the significance level considered is 5%.

TABLE 3. Original model (2 dimensions).

Variable	Factorial weight	SE	*p*-value	R²
1	.880	.016	0	.774
2	.854	.019	0	.729
3	.783	.017	0	.614
4	.806	.018	0	.649
5	.732	.021	0	.536
6	.818	.018	0	.670
7	.814	.018	0	.663
8	.612	.025	0	.374
9	.719	.021	0	.517
10	.780	.018	0	.609
11	.838	.016	0	.703
12	-.817	.024	0	.668
13	.739	.019	0	.546
14	.858	.015	0	.735
15	.471	.035	0	.222
16	.452	.033	0	.204
17	.692	.029	0	.479
18	.614	.032	0	.377
19	.569	.034	0	.324
20	.653	.031	0	.426
21	**-.099**	**.039**	**.01**	**.010**
22	.501	.034	0	.251
23	.579	.032	0	.336
24	**.131**	**.041**	**.001**	**.017**
25	.680	.026	0	.463
26	.510	.032	0	.260
27	.551	.031	0	.303

Based on the modification indices, it was observed that item 12 better fitted the *interest in use* (IU) dimension, and so the decision was taken to move it to this dimension. The analysis of the modification indices showed the existence of correlations between the errors from items 22 and 23, items 7 and 8, and items 26 and 27, which were included. As Table 4 shows, with the changes made to the model with two dimensions, a notable but insufficient improvement in the fit values was obtained.

Next, various models were tested that matched the different theoretical assumptions explored: 1 dimension, 4 dimensions, 5 dimensions, 6 dimensions, and 5 and 6 dimension models with second-order factors (see Tables 1 and 4).

As for the fit of these models, it was found that items 19 and 22 always displayed inadequate R^2 values and that the modification indices displayed the same change in item 12. For this reason, all of the models include these modifications.

TABLE 4. Comparison of all of the models studied

	X^2	df	p	X^2/df	**RMSEA**	**RMSEA** (p)	**CGI**	**TLI**
2 factors (original)	2493	323	0	7.718	.093	0	.837	.822
2 factors	1274.986	271	0	4.705	.069	0	.923	.914
1 factor	3631.954	272	0	13.353	.126	0	.741	.714
4 factors	1208.161	267	0	4.525	.067	0	.927	.918
5 factors	974.259	263	0	3.704	.059	0	.945	.937
5 factors + 2 second order factors	1028.209	268	0	3.837	.060	0	.941	.935
6 factors	965.907	258	0	3.744	.059	0	.945	.937
6 factors + 2 second order factors	1113.053	267	0	4.169	.064	0	.935	.927

On this point, it should be noted that, in all of the models, the correlations between the factors were defined as free parameters. To analyse the discriminant validity of the model, Table 4 shown a comparison of the indices of fit for the different models tested. The calculated indices of fit show that the one-dimensional structure of the questionnaire should be rejected. The values of χ^2/df, RMSEA, CFI, and TLI do not indicate a clear improvement in fit when we move from the 2-dimension model (IU and PD) to the one with 4 dimensions (IU, CU, OL, and A). However, a clear improvement in fit is observed when comparing the 2-dimension model with the 5-dimension model (PU, BFU, CU, OL, A). With this latter model, a RMSEA value very close to .05 is obtained as well as CFI and TLI values greater than .93, something that indicates an adequate fit of the

model. The 6-dimension model (PUConc, PUComp, BFU, CU, OL, A) also has adequate fit; the theoretical justification is that the 5 and 6 dimension models only differ in the subdivision of the PU dimension into PUConc and PUComp. However, there is no improvement in the indices of fit compared to the 5-dimension model. Finally, with regards to models with 5 and 6 dimensions with second-order factors, there is a small improvement in the incremental fit indices but the absolute fit indices become slightly worse and so no relevant overall improvement is observed. Therefore, the analysis of discriminant validity and the principle of parsimony (Carroll, 1978; Ferrando & Anguiano-Carrasco, 2010) indicate that the model which is simplest and offers

the best fit to the theoretical principles is the five factor model: *perceived usefulness* (PU), *beliefs and feelings about the use of VS* (BFU), *complexity of use* (CU), *obstacles for learning* (OL), and *availability* (A). In this model, the two initial dimensions of IU and PD have been subdivided into two and three, respectively. In this way, more concrete dimensions are obtained that provide valuable information about the aspects that affect teachers' attitude towards the use of simulations (Table 5). Figure 1 shows the outline of the final five factor model of the CADUSV questionnaire, in which a correlation was established between the errors from items 7 and 8 and items 22 and 23 based on what is shown by the modification indices of the model.

TABLE 5. Five factor model.

	Factorial weight	SE	*p*-value	R²
Perceived usefulness				
Item 1	.899	.015	0	.808
Item 2	.876	.018	0	.767
Item 3	.802	.017	0	.643
Item 4	.824	.018	0	.679
Item 5	.746	.021	0	.557
Item 6	.838	.017	0	.702
Item 7	.813	.018	0	.661
Item 8	.602	.026	0	.362
Item 9	.734	.02	0	.539
Beliefs and feelings about the use of virtual simulations				
Item 10	.804	.018	0	.646
Item 11	.865	.016	0	.748

Item 12	.731	.021	0	.534
Item 13	.775	.017	0	.601
Item 14	.891	.014	0	.794
Complexity of use				
Item 15	.59	.033	0	.348
Item 16	.558	.034	0	.311
Item 17	.84	.028	0	.706
Item 18	.742	.03	0	.551
Obstacles for learning				
Item 19	.6	.033	0	.360
Item 20	.692	.032	0	.479
Item 22	.5	.035	0	.250
Item 23	.58	.033	0	.336
Item 25	.701	.029	0	.491
Availability				
Item 26	.876	.021	0	.767
Item 27	.956	.022	0	.914

FIGURE 1. Summary outline of the final 5-factor model.
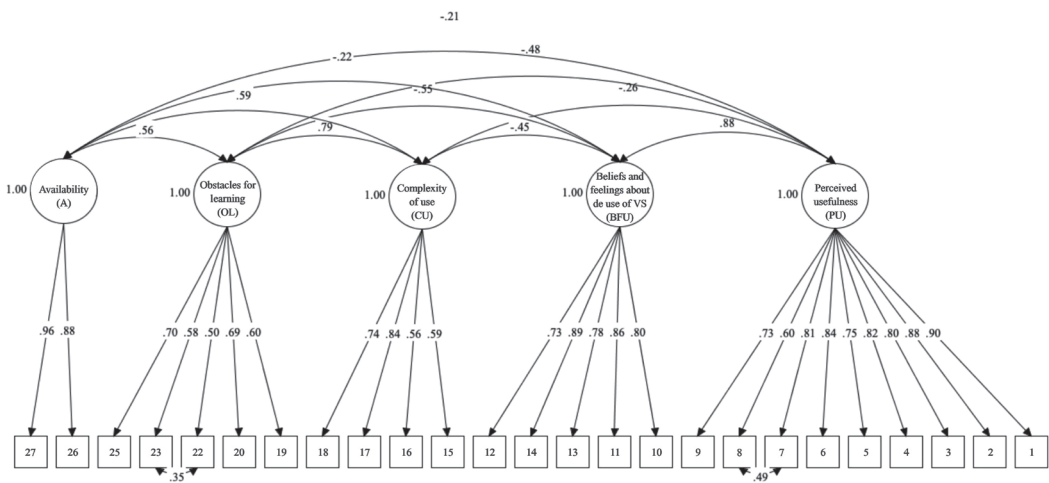
To determine the convergent validity of the 5-factor model, it should be noted that the factor loadings of all of the items are greater than .5 (Table 5). Furthermore, Table 6 shows the results for composite reliability and average variance extracted for the different models, which confirm that the 5-factor model is the most satisfactory in overall terms. The PU, BFU, and A factors display an omega reliability greater than .9 and a mean explained variance much greater than .5. The CU factor displays an omega reliability greater than .7 and an AVE very close to .5 (.479). Finally, the OL factor also displays an omega reliability greater than .7, but an AVE on the limit of what is considered to be acceptable (.37-.5) according to Moral (2019).

TABLE 6. Convergent validity.

Number of factors	Name of factor (initials)	AVE	Omega reliability
1 factor	Attitudes towards use (AU)	.385	.734
2 factors	Interest in use (IU)	.607	.955
	Perceived difficulty (PD)	.335	.854
4 factors	Interest in use (IU)	.607	.955
	Complexity of use (CU)	.478	.781
	Obstacles for learning (OL)	.384	.754
	Availability (A)	.840	.913
5 factors	Perceived usefulness (PU)	.635	.939
	Beliefs and feelings about the use of VS (BFU)	.665	.908
	Complexity of use (CU)	.479	.781
	Obstacles for learning (OL)	.383	.754
	Availability (A)	.841	.913
5 factors + 2 second order factors	Perceived usefulness (PU)	.635	.939
	Beliefs and feelings about the use of VS (BFU)	.665	.908
	Complexity of use (CU)	.479	.782
	Obstacles for learning (OL)	.385	.754
	Availability (A)	.840	.913
	F1: interest in use (IU)	.890	.942
	F2: perceived difficulty (PD)	.666	.851

6 factors	Perceived usefulness for conceptual learning (PUConc)	.698	.933
	Perceived usefulness for competence learning (PUComp)	.618	.827
	Beliefs and feelings about the use of VS (BFU)	.665	.908
	Complexity of use (CU)	.479	.781
	Obstacles for learning (OL)	.384	.754
	Availability (A)	.841	.913
6 factors + 2 second order factors	Perceived usefulness for conceptual learning (PUConc)	.698	.933
	Perceived usefulness for competence learning (PUComp)	.711	.827
	Beliefs and feelings about the use of VS (BFU)	.664	.908
	Complexity of use (CU)	.479	.782
	Obstacles for learning (OL)	.385	.754
	Availability (A)	.841	.913
	F1: interest in use (IU)	.859	.948
	F2: perceived difficulty (PD)	.666	.851

4. Discussion and conclusions

In the present research, a specific questionnaire on teachers' beliefs and attitudes towards the use of virtual simulations was designed and validated to consider in greater depth the main aspects that encourage or discourage teachers from using this tool.

The design of the questionnaire started from the principal factors that influence the use of technologies, such as their usefulness and the ease or complexity of their use (Davis et al., 1989). In addition to these factors, variables described as shaping the behaviour of the teacher were introduced, both in the general use of technologies (Albirini, 2006; Sahin et al., 2016; Teo, 2008; Teo et al., 2016), and in the specific use of virtual simulations (Lee et al., 2021; Zacharia, 2003).

The questionnaire, which initially comprised 46 items, was optimised following the expert judgement and the pilot trial giving a final composition of 27 items. Following the CFA carried out in this work, which starts form an initial two-factor model (*interest in use* and *perceived difficulty*), a scale comprising five

factors was obtained: *perceived usefulness* (PU), comprising 9 items; *beliefs and feelings* towards the use of SV (BFU), comprising 5 items; *complexity of use* (CU), which includes 4 items; *obstacles for learning* (OL), with 5 items; and *availability* (A), which contains 2 items. The division of the *interest in use* theoretical dimension into two factors (PU and BFU) is theoretically based on the differentiation between the aspects referring to the teacher's beliefs regarding the usefulness of virtual simulations (Teo, 2008; Teo et al., 2016) and the effective aspects relating to the emotions of the teacher on the use of SV, previously described by Albirini (2006) or Lee et al. (2021). With regards to the *perceived difficulty* theoretical dimension, this is divided into three factors (CU, OL, and A) because this difficulty, as Lee et al. (2021) previously noted, can have different origins. The results of this work indicate that there are at least three factors that might discourage teachers: complexity of use, previously defined by Teo (2008); availability, closely related to the concept of accessibility studied by Hew and Brush (2007); and obstacles to learning, that is to say, the characteristics described regarding virtual simulations and their use. This five-factor structure provides a greater specificity when analysing teachers' beliefs and attitudes towards the use of virtual simulations and so it can help to identify the decisive factors that encourage or discourage science teachers in the use of virtual simulations in their classrooms.

The quality of the questionnaire has been proven through CFA, providing in-dices of fit that prove the discriminant validity of the 5-factor model. The convergent validity showed satisfactory values for all of the dimensions apart from OL, where despite it having adequate composite reliability values, the AVE values were on the limit of what is acceptable according to Moral (2019), and so it would be interesting to evaluate this dimension to improve its convergent validity in future works. The *availability* (A) dimension displays only two items, and so, despite presenting adequate validity values, it would be interesting to analyse it in greater depth. We consider that these limitations of the questionnaire derive from the very theoretical design of its items. Even though the questionnaires on teachers' attitudes towards the use of technology that we identified were used as a basis, the small number of works published on teachers' attitudes towards the specific use of virtual simulations (Lee et al., 2021; Lehtinen et al., 2016; Zacharia, 2003) and the lack of previous questionnaires in this field hampered the design of the items and dimensions of the questionnaire.

Another limitation of the study relates to the use of non-probability convenience sampling. It would be interesting to administer this questionnaire in future works to a larger sample obtained by means of probability sampling to confirm its validity and improve it.

Overall, the statistical results confirm the validity of the theoretical construct, and so the CADUSV questionnaire appears to be a useful and practical instrument for analysing teachers' attitudes

towards the use of virtual simulations in STEM subjects. Knowledge of this perception is vital to determine the main factors that push teachers away from habitually using this resource, as their perceptions of and attitudes towards technologies directly influence their effective use (Paraskeva et al., 2008). Identifying these factors is also of great interest for designing personalised training proposals that take into account the prior attitudes and beliefs of the teacher (Gargallo, et al., 2006; Padilla, 2018). As Lee et al. (2021) and Pozuelo et al. (2023) already noted, it is necessary to resolve the possible demands or difficulties that teachers face when using virtual simulations to improve their attitude and so increase the effective use of this tool in science teaching.

Authors' contributions

Alicia Palacios: Formal analysis; Investigation; Methodology; Software; Writing (original draft); Writing (review and editing).

Rosa Gómez: Formal analysis; Investigation; Methodology; Software; Writing (original draft); Writing (review and editing).

Álvaro Barreras: Data curation; Investigation; Methodology; Software; Writing (original draft); Writing (review and editing).

Daniel Moreno-Mediavilla: Conceptualization; Investigation; Methodology; Project administration; Writing (originaldraft); Writing (review and editing).

References

Albirini, A. (2006). Teachers' attitudes toward information and communication technologies: The case of Syrian EFL teachers. *Computers & Education*, *47* (4), 373–398. https://doi.org/10.1016/j.compedu.2004.10.013

Amin, D. I., & Ikhsan, J. (2021). Improving higher order thinking skills via semi second life. *European Journal of Educational Research*, *10* (1), 261-274. https://doi.org/10.12973/eu-jer.10.1.261

Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, *105*, 1-13. https://doi.org/10.1016/j.compedu.2016.11.003

Carroll, J. B. (1978). How shall we study individual differences in cognitive abilities? Methodological and theoretical perspectives. *Intelligence*, *2* (2), 87-115. https://doi.org/10.1016/0160-2896(78)90002-8

Chan, P, Van Gerven, T., Dubois, J., & Bernaerts, K. (2021). Virtual chemical laboratories: A systematic literature review of research, technologies and instructional design. *Computers and Education Open*, *2*, 100053. https://doi.org/10.1016/j.caeo.2021.100053

Chen, S., Chang, W., Lai, C., & Tsai, C. (2014). A comparison of students' approaches to inquiry, conceptual learning, and attitudes in simulation-based and microcomputer-based laboratories. *Science Education*, *98* (5), 905-935. https://doi.org/10.1002/sce.21126

D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014*). Simulations for STEM learning: Systematic review and meta-analysis*. SRI Ecucation.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, *35* (8), 982-1003. https://doi.org/10.1287/mnsc.35.8.982

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68* (2), 179-201. https://doi.org/10.3102/00346543068002179

de Pro, A., de Pro, C., & Cantó, J. (2022). Cinco problemas en la formación de maestros y maestras para enseñar ciencias en Educación Primaria. *Revista Interuniversitaria de Formación del Profesorado*, *97* (36.1), 185-202. https://doi.org/10.47553/rifop.v97i36.1.92510

Durán, M. J., Gallardo, S., Toral, S. L., Martínez-Torres, R., & Barrero, F. J. (2007). A learning methodology using Matlab/Simulink for undergraduate electrical engineering courses attending to learner satisfaction outcomes. *International Journal of Technology and Design Education*, *17* (1), 55-73. https://doi.org/10.1007/s10798-006-9007-z

Fan, X., & Geelan, D. (2013). Enhancing students' scientific literacy in science education using interactive simulations: A critical literature review. *Journal of Computers in Mathematics and Science Teaching*, *32* (2), 125-171.

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología [Factor analysis as a research technique in psychology]. *Papeles del Psicólogo*, *31* (1), 18-3. https://www.papelesdelpsicologo.es/pdf/1793.pdf

Gargallo, B., Suárez, J., & Almerich, G. (2006). La influencia de las actitudes de los profesores en el uso de las nuevas tecnologías. **Revista Española de Pedagogía**, *64* (233), 45-66.

Gómez, R., Palacios, A., Moreno-Mediavilla, D., & Barreras, Á. (2022). Competencias docentes en el uso de simulaciones virtuales STEM: diseño y validación de un instrumento de medida (CDUSV) [Teacher competences in the use of STEM virtual simulations: design and validation of a measurement instrument (CDUSV)]. Bordón, *Revista de Pedagogía, 74* (4), 85-102.

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing y Health*, *20* (3), 269-274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3%3C269::AID-NUR9%3E3.0.CO;2-G

Green, S. B., & Yang, Y. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: Coefficient alpha and omega coefficients. *Educational Measurement: Issues and Practices*, *34* (4), 14-20. https://doi.org/10.1111/emip.12100

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8.ª ed.). Cengage Learning.

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research and Development*, *55* (3), 223-252. https://doi.org/10.1007/s11423-006-9022-5

Khan, S. (2011). New pedagogies on teaching science with computer simulations. *Journal of Science Education and Technology*, *20* (3), 215-232. https://doi.org/10.1007/s10956-010-9247-2

Lai, C., Wang, Q., & Lei, J. (2012). What factors predict undergraduate students' use of technology for learning? A case from Hong Kong. *Computers & Education*, *59* (2), 569-579. https://doi.org/10.1016/j.compedu.2012.03.006

Lee, W. C., Neo, W. L., Chen, D. T., & Lin, T. B. (2021). Fostering changes in teacher attitudes toward the use of computer simulations: Flexibility, pedagogy, usability and needs. *Education and Information Technologies*, *26* (4), 4905-4923. https://doi.org/10.1007/s10639-021-10506-2

Lehtinen, A., Nieminen, P., & Viiri, J. (2016). Preservice teachers' TPACK beliefs and attitudes toward simulations. *Contemporary Issues in Technology and Teacher Education*, *16* (2), 151-171.

Lynch, T., & Ghergulescu, I. (2017). Review of virtual labs as the emerging technologies for teaching STEM subjects. En L. Gómez, A. López, & I. Candel (Eds.), *INTED2017 Proceedings. 11th International Technology, Education and Development Conference. March 6th-8th, 2017 - Valencia, Spain* (pp. 6082-6091). IATED Academy. https://doi.org/10.21125/inted.2017.1422

Mardia, K. V. (1970). Measures of multivariate skewnees and kurtosis with applications. *Biometrika*, *57* (3), 519-530. https://doi.org/10.2307/2334770

Moral, J. (2019). Revisión de los criterios para validez convergente estimada a través de la varianza media extraída [Review of the criteria for convergent validity estimated through the extracted average variance]. *Psychología*, *13* (2), 25-41. https://revistas.usb.edu.co/index.php/Psychologia/article/view/4119

Moreno-Mediavilla, D., Palacios-Ortega, A., Gómez, R., & Barreras-Peral, A. (2023). Competencia digital docente en el uso de simulaciones virtuales: percepción del profesorado de áreas STEM [Teacher digital competence in the use of virtual simulations: STEM teachers' perception]. *Pixel-Bit. Revista de Medios y Educación*, (68), 83-113. https://doi.org/10.12795/pixelbit.98768

Nunes, P. S., Nasciminto, M. M., Catarino, P., & Martins, P. (2020). Fatores que influenciam o uso de software educativo no ensino de matemática [Factors that influence the use of educational software in mathematics teaching]. *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, *18* (3), 113-129. https://doi.org/10.15366/reice2020.18.3.006

Oliveira, A., Feyzi, R., Behnagh, Ni, L., Mohsinah, A., Burguess, K., & Guo, L. (2019). Emerging technologies as pedagogical tools for teaching and learning science: A literature review. *Human Behaviour and Emerging Technology*, *1* (2), 149-160. https://doi.org/10.1002/hbe2.141

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25* (2), 177-196. https://doi.org/10.1007/s10972-014-9384-1

Padilla, S. (2018). Usos y actitudes de los formadores de docentes ante las TIC. Entre lo recomendable y la realidad de las aulas [Teachers trainer's uses and attitudes before ICT. Between the advisible and the reality of classrooms]. *Apertura*, *10* (1), 132-148. https://doi.org/10.32870/ap.v10n1.1107

Paraskeva, F., Bouta, H., & Papagianni, A. (2008). Individual characteristics and computer self-efficacy in secondary education teachers to integrate technology in educational practice. *Computers y Education*, *50* (3), 1084-1091. https://doi.org/10.1016/j.compedu.2006.10.006

Plass, J. L., Milne, C., Homer, B. D., Schwartz, R. N., Hayward, E. O., Jordan, T., Verkuilen, J., Ng, F., Wang, Y., & Barrientos, J. (2012). Investigating the effectiveness of computer simulations for chemistry learning. *Journal of Research in Science Teaching*, *49* (3), 394-419. https://doi.org/10.1002/tea.21008

Pozuelo, J., Martín, J., Carrasquer, B., & Cascarosa, E. (2023). Percepciones del profesorado ante el uso de simuladores virtuales en el aula de ciencias [Teachers' perceptions about the use of virtual simulators in the science classroom]. *Revista Interuniversitaria de Formación del Profesorado. Continuación de la Antigua Revista de Escuelas Normales*, *98* (37.2). https://doi.org/10.47553/rifop.v98i37.2.95842

Rutten, N., Van Joolingen, W. R., & Van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers and Education*, *58* (1), 136-153. https://doi.org/10.1016/j.compedu.2011.07.017

Sahin, A., Top, N., & Delen, E. (2016). Teachers' first-year experience with chromebook laptops and their attitudes towards technology integration. *Technology, knowledge, and learning*, *21* (3), 361-378. https://doi.org/10.1007/s10758-016-9277-9

Shapka, J. D., & Ferrari, M. (2003). Computer-related attitudes and actions of teacher candidates. *Computers in Human Behavior, 19* (3), 319-334. https://doi.org/10.1016/S0747-5632(02)00059-6

Stieff, M. (2019). Improving learning outcomes in secondary chemistry with visualization-supported inquiry activities. *Journal of Chemical Education*, *96* (7), 1300-1307. https://doi.org/10.1021/acs.jchemed.9b00205

Tate, M., Evermann, J., & Gable, G. (2015). An integrated framework for theories of individual attitudes toward technology. *Information y Management*, *52* (6), 710-727. https://doi.org/10.1016/j.im.2015.06.005

Teo, T. (2008). Pre-service teachers' attitudes towards computer use: A Singapore survey. *Australasian Journal of Educational Technology*, *24* (4), 413-424. https://doi.org/10.14742/ajet.1201

Teo, T. (2012). Examining the intention to use technology among pre-service teachers: An integration of the technology acceptance model (TAM) and theory of planned behavior (TPB). *Interactive Learning Environments*, *20* (1), 3-18. https://doi.org/10.1080/10494821003714632

Teo, T., Milutinović, V., & Zhou, M. (2016). Modelling Serbian pre-service teachers' attitudes towards computer use: A SEM and MIMIC approach. *Computers* y *Education*, *94*, 77-88. https://doi.org/10.1016/j.compedu.2015.10.022

Trujillo, W., Curo, L., Paredes, L., & Carbajal, K. (2023). Eficiencia de los simuladores virtuales en la competencia de indagación para el aprendizaje de física elemental [Efficiency of virtual simulators in the inquiry competition for learning elementary physics]. *Telos: Revista de Estudios Interdisciplinarios en Ciencias Sociales*, *25* (2), 459-476.

Waight, N., Liu, X., Gregorius, R. M., Smith, E., & Park M. (2014) Teacher conceptions and approaches associated with an immersive instructional implementation of computer-based models and assessment in a secondary chemistry classroom. *International Journal of Science Education*, *36* (3), 467-505, https://doi.org/10.1080/09500693.2013.787506

Wen, C., Liu, C., Chang, H., Chang, C., Chang, M., Chiang, S. F., Yang, C., & Hwang, F. (2020). Students' guided inquiry with simulation and its relation to school science achievement and scientific literacy. *Computers* y *Education*, *149*. https://doi.org/10.1016/j.compedu.2020.103830

Wong, K. T., Teo, T. y Russo, S. (2012). Influence of gender and computer teaching efficacy on computer acceptance among Malaysian student teachers: An extended Technology Acceptance Model (TAM). *Australasian Journal of Educational Technology*, *28* (7), 1190-120. https://doi.org/10.14742/ajet.796

Wu, H. K., & Huang, Y. L. (2007). Ninth-grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. *Science Education, 91*(5), 727-749. https://doi.org/10.1002/sce.20216

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Method*, *51*, 409 - 428. https://doi.org/10.3758/s13428-018-1055-2

Zacharia, Z. (2003). Beliefs, attitudes, and intentions of science teachers regarding the educational use of computer simulations and inquiry-based experiments in physics. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *40* (8), 792-823. https://doi.org/10.1002/tea.10112

Authors' biographies

Alicia Palacios. Doctor of Biochemistry and Molecular Biology from the Universidad Complutense de Madrid. Accredited as an associate professor by ANECA. Academic Director of the Master's in Didactics of Physics and Chemistry at UNIR and professor of didactics of the experimental sciences in the Faculty of Education. Broad experience in the development of teaching innovation projects. Secondary education science teacher and editor for the publisher Edebé. Her research activity centres on the study of didactic resources and methodologies that foster active and contextualised learning in the sciences She is part of the "Didactics of Mathematics and the Experimental Sciences" (DIMACE) research group.

https://orcid.org/0000-0002-7906-1417

Rosa Gómez. Doctor of Research in Teaching and Learning of Experimental and Social Sciences and Mathematics from the Universidad de Extremadura. Professor and researcher in the Faculty of Education at the Universidad Internacional de La Rioja (UNIR) since 2017. Member of the "Didactics of Mathematics and the Experimental Sciences" (DIMACE) research group.

https://orcid.org/0000-0001-5861-9429

Álvaro Barreras. Doctor of Mathematics. Professor of mathematics (Centro Universitario de la Defensa de Zaragoza, 2014–2016). Professor of didactics of mathematics (UNIR, 2016–present). Academic director of two didactics of mathematics master's programmes. Lead researchers in the "Didactics of Mathematics and the Experimental Sciences" (DIMACE) research group at the UNIR.

 https://orcid.org/0000-0001-5325-8505

Daniel Moreno-Mediavilla. Professor in the Faculty of Education of the Universidad Internacional de La Rioja (UNIR). Member of the "Didactics of Mathematics and the Experimental Sciences" (DIMACE) research group. Doctor of Chemistry from the Universidad de Burgos. More than 20 articles published in JCR and Scopus, and author of a patent.

 https://orcid.org/0000-0002-5633-2376

Chilean teachers' cultural threat and attitudes towards multiculturality at school: The role of outgroup anxiety and intercultural sensitivity

La amenaza cultural y la actitud de los docentes chilenos hacia la multiculturalidad en la escuela: el rol de la ansiedad exogrupal y la sensibilidad intercultural

María-José MERA-LEMP, PhD. Scholar. School of Law and Social Sciences, Universidad Viña del Mar, Chile (*mariajosemera@yahoo.es*).

Javier TORRES-VALLEJOS, PhD. Scholar. School of Psychology, Faculty of Social Sciences and Communications, Universidad Santo Tomas, Chile (*jtorresvallejos@santotomas.cl*).

Florencia GUGLIELMETTI-SERRANO, MPsych. Scholar. School of Psychology, Universidad Gabriela Mistral, Chile (*fguglielmettis@gmail.com*).

Abstract:

In Chile, the arrival of immigrant students has had significant consequences for teachers, who have had to adapt quickly and with few resources to a new school reality. This makes it relevant to study variables affecting their attitudes towards immigrant students. This work aimed to determine the role of outgroup anxiety and intercultural sensitivity in the relationship between cultural threat and attitudes toward multiculturalism at school in a sample of 190 primary and secondary education teachers. A sequential mediation model was carried out. Outcomes show the direct effect of cultural threat on teachers' attitudes towards multiculturalism at school, as well as its indirect effects through outgroup anxiety and intercultural sensitivity. The model explains 55% of the variability of the dependent variable. Results stress the importance of attending to the socio-cognitive and emotional processes affecting teachers' beliefs and attitudes regarding immigrant students to facilitate teaching work and school relationships.

Keywords: teachers, attitudes toward multi-culturality at school, cultural threat, outgroup anxiety, intercultural sensitivity.

Resumen:

En Chile, la llegada de escolares inmigrantes ha tenido importantes consecuencias para el profesorado, que ha debido adaptarse de manera rápida y con escasos recursos a una nueva realidad en las aulas. Esto pone de relevancia el estudio de variables que inciden en las actitudes que desarrollan hacia el alumnado inmigrante. Este trabajo tuvo como objetivo determinar el rol de la ansiedad exogrupal y la sensibilidad intercultural en la relación entre la amenaza cultural y la actitud hacia la multiculturalidad en la escuela, en una muestra compuesta por 190 profesores de educación primaria y secundaria. Se realizó un modelo de mediación secuencial que muestra el efecto directo de la amenaza cultural en la actitud de los docentes hacia la multiculturalidad en la escuela, así como sus efectos indirectos a través de la ansiedad exo-grupal y la sensibilidad intercultural. El modelo explica el 55 % de la variabilidad de la variable dependiente. Se discute acerca de la necesidad de atender a los procesos sociocognitivos y emocionales que inciden en las creencias y actitudes de los profesores respecto del alumnado inmigrante con el fin de facilitar la labor docente y la convivencia escolar.

Palabras clave: profesores, actitudes hacia la multiculturalidad, amenaza cultural, ansiedad exogrupal, sensibilidad intercultural.

1. Introduction

The Chilean education system faces relevant challenges due to the fast and substantial increase of immigrant students (Ministerio de Educación, Centro de Estudios, 2023). During the last few years, intraregional migration in Latin America and the Caribbean has increased considerably, with Chile becoming one of the main reception countries (Organización Internacional del Trabajo, 2023).

By 2022, foreigners represented 8.3% of the total inhabitants of the country, hailing from countries such as Venezuela (32.8%), Peru (15.4%), Colombia (11.7%) and Haiti (11.4%) (Instituto Nacional de Estadísticas y Departamento de Extranjería y Migración, 2022). One of the main characteristics of these new migration flows is the growing presence of foreign school-aged children and adolescents, which resulted in a 15.7% increase in school enrolment of immigrant students between 2017 and 2022, representing 5.3% of total national school enrolments (Ministerio de Educación, Centro de Estudios, 2023).

For Chilean teachers, teaching in this new multicultural context means confronting not only pedagogical challenges but also the relationship with immigrant students (Flanagan-Bórquez et al., 2021). These teachers have been trained with monocultural curriculums where intercultural education has a

very minor role and focuses on indigenous groups (Riedemann et al., 2020). Furthermore, even though the Chilean State has made efforts to include the intercultural perspective in education policies, major difficulties remain in its implementation (González et al., 2023; Mera-Lemp et al., 2021; Ortega et al., 2020).

Several studies have suggested Chilean teachers' attitudes towards multiculturalism in school tend to be ambivalent and negative. The causes would be related to racism (Tijoux & Zapata-Sepúlveda, 2019), as well as to the belief that immigrant students possess educational delays and behaviors that are incompatible with the ones expected by schools (Cerón et al., 2017; Pavez-Soto et al., 2019b; Salas et al., 2017; Stefoni & Corvalán, 2019). This would translate into teaching practices that demand the abandonment of the students' own cultural identities in the class environment to the detriment of their integration into the school community (Jiménez & Fardella, 2015).

1.1. Teachers' attitudes towards multiculturality within the school

There is consensus on how attitude is defined as the evaluative response that people have towards members of social groups they do not belong to, which entails a cognitive, affective, and behavioral component (Allport, 1968).

Specifically, attitudes towards multiculturality or multiculturalism have been defined as the positive valuation of cultural diversity, the support for the participation of minority groups in all aspects of society, and the acceptance that all citizens, including members of cultural majorities, must adapt to one another (Berry et al., 1977). Based upon these definitions, attitudes towards multiculturalism in school reflect how teachers evaluate the presence of immigrant students in their centers and its possible consequences (León et al., 2007).

Research has suggested that the work of teachers can be taut when there are difficulties in positively recognizing cultural diversity and including it as a resource in the daily work with students (Gutentag et al., 2018). Teaching has been characterized as an emotionally high-demand job, usually accompanied by high levels of stress and frustration (Fisher, 2011). Numerous studies (Dubbeld et al., 2017; Glock et al., 2019; Gutentag et al., 2018) have shown that teaching immigrant students can become a stressor and diminish teachers' self-efficacy when facing difficulties implementing pertinent and effective pedagogical strategies.

Furthermore, teachers are key agents for students' socialization and positive development. The quality of teachers' relationships with their students has important consequences for self-esteem, well-being, and cultural adjustment to the reception country in the case of those who are immigrants (Makarova et al., 2019; Valcke et al., 2022). In addition, through their

attitudes and behaviors in classrooms, teachers transmit values, beliefs, and ways of socializing with people who are different, which affects the relationships between immigrant and local students, promoting their integration or, on the contrary, influencing the existence of conflicts (Grütter et al., 2021; Horenczyk & Tatar, 2002). These highlight the importance of studying the attitudes of teachers towards multiculturality in schools due to the implications on their well-being, teaching practices, as well as their impact on the development and socialization of their students (Leix, 2015; Zanga & De Gioannis, 2023).

1.2. The effect of cultural threat, outgroup anxiety, and intercultural sensitivity on teachers' attitudes toward multiculturality

Literature suggests that the degree to which teachers positively evaluate cultural diversity caused by the incorporation of immigrant students depends in great measure on how they perceive these groups as carriers of values, beliefs, and behavioral cues that go against the culture of the societal majority (Mera-Lemp et al., 2021; Nguyen & Boreczky, 2022).

According to the intergroup threat theory (Stephan et al., 2015), the perception of immigrant groups' culture as a threat would cause negative emotions in teachers. These emotions, which have been conceptualized as *outgroup anxiety*, would activate attitudes, stereotypes, and negative beliefs about the outgroup. Intergroup anxiety is related to fear, anger, frustration, and guilt, among other emotions (Stephan, 2014), which would predispose people to react negatively to outgroup members, affecting nonverbal language (Mendes & Koslov, 2012; Trawalter et al., 2012) and negatively impacting effective intercultural communication (Ulrey & Amason, 2001).

Through this mechanism, the anxiety provoked by the perception of threat would affect the intercultural sensitivity of teachers, understanding it as the ability to successfully communicate in cultural diversity contexts (Chen & Starosta, 2000; Khukhlaev et al., 2022). Intercultural sensitivity means attending to and respecting cultural differences in communicative interactions, trusting in one's capacities to establish these relationships, and, consequently, involving oneself and enjoying the interaction (Chen & Starosta, 2000). This competency would positively influence teachers' attitudes towards multiculturality in school, which could propitiate the development of culturally sensitive and inclusive pedagogical practices (Akcaoğlu & Kayis, 2021; Argüello-Gutiérrez et al., 2024; Blazevic et al., 2023; Kim & Connelly, 2019).

From this background, this work aimed to determine the mediating role of outgroup anxiety and intercultural sensitivity in the relationship between cultural threat and attitudes towards multiculturality at school in a sample of Chilean teachers.

On the one hand, it is hypothesized that outgroup anxiety will present a negative mediating effect in the relationship between cultural threat and the attitude towards multiculturality at school. On the other hand, as a second hypothesis, intercultural sensitivity will have a positive mediating effect. Despite the evidence in favor of the association between the dependent and independent variables, we have not found studies that analyze what variables can intervene in this relationship and consider intercultural sensitivity.

2. Method

2.1. Participants

The study used a non-probabilistic convenience sampling method. The inclusion criteria considered that the participants were Chilean; with bachelor's degrees in education; working in public, private, and schools financed by the State but administrated by private organizations; and working at the primary or secondary level. Foreign teachers and those who did not have a bachelor's degree in education or worked at special education or learning leveling centers were excluded from the sample.

The sample comprised 190 Chilean teachers (70% female, 30% male) between the ages of 24 and 65 (M = 40.48; SD = 11.22). More than half of the teachers (58.8%) worked at the primary education level, while the rest (41.1%) worked at the secondary education level. Moreover, 47.9% of the sample

taught at public schools, 27.9% worked at schools financed by the State but administrated by private organizations, and 24.2% worked at private schools. 51.2% of the sample worked in classrooms with immigrant students. According to their family income, 43.2% of the teachers had a low socioeconomic level, 44.7% had a middle socioeconomic level, and 12.1% had a high socioeconomic level.

2.2. Variables and instruments

Attitude toward multiculturality at school. The attitude towards multiculturality at school scale (León et al., 2007), validated in Chile by Mera-Lemp et al. (2021), was applied. It consists of 8 Likert-type items, with five answer options (1 = Completely disagree, 5 = Completely agree); e. g., "The presence of immigrant students at school negatively affects group academic performance". The instrument's reliability in this study was α = .74 and Ω = .86. The model fit was CFI = .942, TLI = .904, RMSEA = .115, SRMR = .064.

Cultural threat. 3 items from the questionnaire of perceptions and attitudes towards foreign population (Aierdi et al., 2004) were used (e. g., "Latin-American immigrants enrich the culture of our region"). These items have five answer options (1 = Completely disagree, 5 = Completely agree). The scale's reliability in this study was α = .80 and Ω = .80. The fit of the instrument was excellent: CFI = .999, TLI = .999, RMSEA = .001, SRMR = .001.

Outgroup anxiety. Koc and Anderson's (2018) adaptation of the original scale constructed by Stephan and Stephan (1985) was applied. The scale consists of 6 Likert-type items with five answer options (1 = Nothing, 5 = A lot). These items asked participants about their emotions when interacting with outgroup members (such as tension, insecurity, and discomfort). This instrument presented a reliability of α = .89 and Ω = .88, and its fit was good: CFI = .993, TLI = .987, RMSEA = .054, SRMR = .021.

Intercultural sensitivity. The intercultural sensitivity scale (Chen & Starosta, 2000), validated in the Chilean population by Martínez-Zelaya et al. (2020), was applied. It consists of 24 Likert-type items with four answer options (1 = Very disagree, 4 = Very agree) (e.g., "I avoid situations in which I am forced to interact with people with different cultural backgrounds"). Scale's reliability in this study was α = .84 and Ω = .85, and the model fit was adequate (CFI = .952, TLI = .933, RMSEA = .045, SRMR = .079).

2.3. Procedure

First, school principals were requested to participate, and then teachers were asked to participate voluntarily. The researchers administered the instruments at the educational centers. The project was approved by the Ethics Committee of Universidad Alberto Hurtado on May 27, 2022, and the necessary ethical procedures were followed to respect human rights in the study involving human subjects.

Teachers signed informed consent forms that guaranteed the confidentiality and anonymity of the produced information. The information was treated confidentially, and any elements that could identify the participants were removed.

2.4. Analytic plan

The data were analyzed using IBM SPSS 24 statistical software. After the reliability analyses of the instruments were conducted, Pearson correlation analyses were performed. The PROCESS macro was used to carry out a sequential mediation model.

3. Results

3.1. Descriptive analysis

First, descriptive analyses were conducted. As shown in Table 1, participants reported positive attitudes toward multiculturalism in schools, with scores above the midpoint on the 5-point scale. Symbolic threat and intergroup anxiety were low, with both variables having means below the midpoint of their respective scales. Regarding intercultural sensitivity, scores were above the midpoint on the 4-point scale. Overall, these results indicate that teachers tend to value the presence of immigrant students in schools positively. Their appreciation of different cultures tends to be positive, and they exhibit low levels of negative emotions in contact with these students. Additionally, these results suggest that their competence in successfully communicating with members of other cultures is adequate.

TABLE 1. Descriptive statistics and bivariate correlations among sociodemographic variables, cultural threat, outgroup anxiety, intercultural sensitivity, and attitude toward multiculturality at school.

	1	2	3	4	5	6	7	8
1	-							
2	.154*	-						
3	-.006	.267**	-					
4	.004	.010	.008	-				
5	.072	.162*	.145*	-.234**	-			
6	.128	.241**	.245**	-.137	.673**	-		
7	-.081	-.113	-.038	.274**	-.593**	-.596**	-	
8	-.136	-.232**	-.131	.123	-.681**	-.630**	.556**	-
M					2.15	1.85	3.14	4.49
SD					.85	.58	.38	.56

Note 1: 1 = age, 2 = gender (female = 0, male = 1), 3 = teaching level (primary = 0, secondary = 1), 4 = presence of immigrant students (no = 0, yes = 1), 5 = cultural threat, 6 = outgroup anxiety, 7 = intercultural sensitivity, 8 = attitude toward multiculturality at school.
Note 2: *p <.05, **p <.01, ***p <.001.

3.2. Bivariate correlation analysis

Before testing the hypotheses, Pearson correlation analyses were conducted between sociodemographic and study variables. These relationships were examined to control for their potential effects on the associations among the variables under investigation. The results (see Table 1) indicate the existence of statistically significant relationships between sociodemographic variables and those that are the focus of this study. Gender was significantly and negatively associated with the attitude toward multiculturalism in schools, suggesting that being female is linked to a more favorable disposition toward the participation of immigrant students in schools.

In contrast, there is a positive relationship between gender, symbolic threat, and intergroup anxiety. This suggests that being male is associated with a greater perception that the cultures of immigrant people may threaten the dominant Chilean culture, as well as

an increase in negative emotions during contact.

Additionally, working at the secondary education level was significantly and positively associated with symbolic threat and intergroup anxiety. This could indicate that secondary school teachers perceive, to a greater extent than primary school teachers, that the culture of immigrant students may negatively affect Chilean culture, and would be linked to greater negative emotions toward them.

Similarly, teaching in classrooms with immigrant students was related to a lower perception of symbolic threat and greater intercultural sensitivity. In other words, those who work with immigrant students tend to hold lower negative beliefs about the culture of these students and show a greater ability to communicate successfully in multicultural contexts than those who do not work with such students. Teacher age did not show statistically significant relationships with any of the variables under study.

Third, statistically significant relationships among the study variables were confirmed. The results show that symbolic threat was significantly and negatively related to the attitude toward multiculturalism in schools and intercultural sensitivity. That is, as the level of perceived threat increases, the positive valuation of the presence of immigrant students decreases, as does the ability to communicate successfully with these students. In contrast, its association with intergroup anxiety was significant and positive, indicating

that, as the perception of threat increases, so do negative emotions during contact with foreign students.

Outgroup anxiety showed significant and negative associations with attitudes toward multiculturalism in schools and with intercultural sensitivity, showing that an increase in negative emotions is associated with a decrease in the acceptance of immigrant students and a lower ability in intercultural communication. The results also reveal statistically significant positive relationships between intercultural sensitivity and attitudes toward multiculturality at school. This indicates that as teachers' ability to communicate with foreign students increases, so does the appreciation of cultural diversity in schools.

3.3. Mediation model

A multiple mediation model was used to test the hypotheses of this study (see Figure 1). Since statistically significant relationships between sociodemographic variables and the study variables had been previously established, these were included in the model as control variables.

The results show that controlling for the effects of gender ($\beta = -.11$, $t = -.911$, $p = .057$), contact with immigrant students ($\beta = -.03$, $t = -.545$, $p = .586$), and the level at which teaching is conducted ($\beta = .01$, $t = .304$, $p = .761$), the cultural threat has a negative and significant effect ($\beta = -.69$, $t = -11.964$, $p = .000$) on the attitude toward multiculturality at school, and this effect decreases when the mediating variables are included (direct effect: $\beta = -.43$, $t = -5.976$, $p = .000$).

FIGURE 1. Sequential mediation model: outgroup anxiety and intercultural sensitivity in the relationship between cultural threat and attitude toward multiculturality at school.

*p <0.05, **p <0.01, ***p <0.001
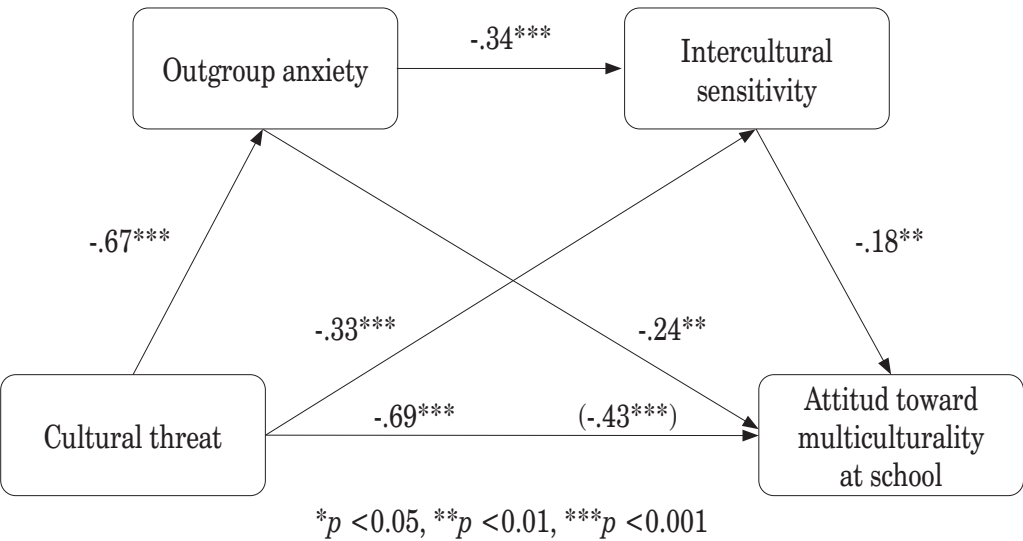
As hypothesized, the indirect effect of cultural threat on attitude toward multiculturality at school through outgroup anxiety is significant (β = -.16, SE = 0.06, 95% CI [-0.2836, -0.0219]). Its indirect effect through intercultural sensitivity is also significant (β = -.05, SE = 0.02, 95% CI [-0.1031, -0.0094]). Furthermore, there is a total sequential indirect effect of cultural threat on attitude toward multiculturality at school through outgroup anxiety and then through intercultural sensitivity (β = -.04, SE = 0.02, 95% CI [-0.0826, -0.0112]). This model accounts for partial mediation and explains 55% of the variability in teachers' attitudes toward multiculturality at school.

4. Discussion

This work aimed to study the incidence of perceived cultural threat on Chilean teachers' attitudes to multiculturality at school,

establishing the role of outgroup anxiety and intercultural sensitivity in this relation.

Even though prior studies conducted with Chilean teachers suggested the existence of negative valuations regarding immigrant people and their participation in school communities (Beniscelli et al., 2019; Salas et al., 2017), our descriptive results revealed that participants in our study tend to show low levels of agreement with the belief that their culture could threat Chilean mainstream culture, and also low levels of anxiety at intergroup contact. Moreover, outcomes point out to a positive valuation of immigrant students' participation at schools. Besides, teachers present high levels of intercultural sensitivity, partially coinciding with former studies conducted with similar samples that have reported positive moderated levels of this competence (Morales et al., 2017). Differences between our results and those

informed by previous research may be due to different methodologies (quantitative versus qualitative) and measuring instruments.

The analyses of the relationships between participants' sociodemographic characteristics and the variables in the study show that female teachers present a better attitude toward multiculturality and lower levels of both threat and anxiety towards immigrant people than male teachers. This coincides with studies that have claimed that male teachers tend to present higher levels of nationalism (McCorkle & Rodríguez, 2023) and that female teachers show fewer levels of negative attitudes toward students from ethnic minorities (Abacioglu et al., 2019; Chin et al., 2020). Meta-analytic studies conducted with general populations have found higher levels of ethnocentrism and prejudice in men than in women, which has been explained by socialization differences (Dozo, 2015). This could be a possible explanation for our results.

Besides, a positive relation between teaching at the secondary level, threat, and anxiety was found. Few studies consider the educational level at which teachers work when studying their attitudes toward immigrant students (Monreal & McClorke, 2020). As a hypothesis, we think that it may occur because teachers working at secondary schools have to deal with students who have been less exposed to mainstream culture and to the host country's educational system, which can represent a great difficulty in carrying out their work. Likewise, secondary school

teachers are usually exposed to stressors produced by conflict management both in classrooms and in their relations with parents, which could be intensified in cultural diversity scenarios (Abós et al., 2019; Fisher, 2011).

The outcomes also show that teachers who work with immigrant students present lower levels of threat and higher intercultural sensitivity than those who have no contact with them. It has been suggested that repeated contact with students of cultural minorities leads to a higher social closeness, facilitating the development of positive attitudes toward them (Chin et al., 2020; Glock et al., 2018; Gutentag et al., 2018; Tatar et al., 2003). Other authors have proposed that schools with greater ethnic diversity would be more likely to generate strategies and provide resources to improve teaching with these students, facilitating the relationship with them and, consequently, contributing to the development of higher intercultural sensitivity in teachers (Alonso, 2011).

The analysis of the relationships between the variables in the study also shows that threat presents a positive association with anxiety, revealing that the perception that immigrants' beliefs, values, and culturally based behaviors can damage the host culture is accompanied by negative emotional experiences at intergroup contact (Stephan & Stephan, 2014).

Conversely, its relationship with intercultural sensitivity was negative. This finding is in line with precedents indicating that the development of this competence

requires overcoming the rejection of cultural differences and ethnocentrism. This would promote respect for differences, confidence that communication with different others will be successful, as well as attentiveness, involvement and enjoyment in interaction (Alexandra, 2018; Chen and Starosta, 1998; Bennet, 1986); thus, better intercultural communication would be fostered.

Likewise, the association between threat and attitude towards multiculturalism at school was also negative, in line with studies that relate it to a lower willingness to integrate different others, as well as a lower degree of adherence to multiculturalism (Cala et al., 2018; Nguyen and Boreczky, 2022). This would be explained by the fact that teachers would feel their cultural identity threatened and would decrease their acceptance of diversity. As a consequence, there would be a greater tendency to exclude migrant schoolchildren and more support for their assimilation into the majority culture (Crozier and Davies, 2008; Mera, et al., 2017).

Outgroup anxiety presents a negative link with intercultural sensitivity, which development depends to a great extent on the emotions perceived towards outgroup members. When intergroup contact leads to negative emotions, people try to avoid this kind of interaction because of the anticipation of failure in establishing positive relationships. Feelings such as insecurity, discomfort, and tension reduce the disposition to respect discrepancies through intercultural encounters (Khukhlaev *et al.*, 2021; Khukhlaev et al., 2022).

Besides, our results show positive relations between intercultural sensitivity and attitudes toward multiculturality at school. This could occur because intercultural competence facilitates the anticipation of positive relationships and a better teaching performance. This is especially important in cultural diversity contexts, which can be perceived as highly demanding, threatening teachers' self-efficacy. Thus, intercultural sensitivity is associated with a favorable disposition towards diversity, promoting a positive teaching climate (Akcaoğlu & Kayis, 2021; Blazevic et al., 2023; Kim & Connelly, 2019).

The sequential mediation model, which explains 55% of the changes on the dependent variable, shows that cultural threat has a direct negative effect on the attitude towards multiculturality at school. This effect decreases in the presence of outgroup anxiety and intercultural sensitivity. As we hypothesized, cultural threat increases anxiety towards immigrants, which, in turn, diminishes intercultural sensitivity, affecting teachers' attitudes towards multiculturality at school.

5. Conclusions

The results highlight the importance of considering the socio-cognitive and emotional processes that intervene in teachers' approaches to multiculturality in the school context. Our findings show the relevant impact of negative beliefs about immigrant students' cultures on teachers' valuations of cultural diversity at schools. Moreover, outcomes reveal that this effect is produced because cultural threat

increases negative emotions towards these students, reducing teachers' capabilities to communicate successfully in intercultural settings.

Teachers carry out the complex task of teaching in culturally diverse contexts. However, they have frequently been trained through education curriculums that reproduce the worldviews of the dominant social groups, with scarce opportunities for developing intercultural competencies. This is especially true in the case of Chilean teachers (Pavez-Soto et al., 2019a; Sánchez-Sánchez et al., 2018; Sanhueza et al., 2014). Literature shows that in these situations, working in multicultural environments could lead to tensions because the application of teaching practices that have been successful with students of their own cultural groups tends to fail when working with minority groups' students. This can diminish teachers' self-efficacy feelings, increasing stress and jeopardizing their wellbeing (Dubbeld et al., 2017; Gutentag et al., 2018).

Besides, their beliefs, expectations, and behaviors have important consequences on minority group students' learning processes and well-being, who are already in a position of psychosocial disadvantage and undergoing processes of psychological, social, and cultural adjustment to the host country (Pastor-Vicedo et al., 2016; Tyler et al., 2006).

Teachers' attitudes towards immigrant students can also have effects on host society students, who vicariously, learn ways of relating to them, facilitating or prevent-

ing intergroup conflict between students (Horenczyk & Tatar, 2002). Thus, teachers' perspectives are critical for the construction of school climates that encourage the positive recognition of differences, justice, and cooperation between school communities' members (Haenni et al., 2019; Ulbricht et al., 2022).

This study presents limitations to be considered. First, the use of an intentional sample in the framework of a cross-sectional design does not allow to appreciate the relationship between variables across time. Second, the sample size does not allow us to study whether there are differences between schools. Third, other variables related to attitudes toward immigrant students were not included, such as acculturation preferences or stereotypes. Also, variables associated with teachers' situations, such as work-related stress, self-efficacy, and perceived school climate, were not included.

Nevertheless, this study contributes to developing a highly important research line due to the increase in multiculturality at educational centers, both in the case of south-south migration and at the global level. This underlines the need to produce knowledge that can guide interventions to facilitate teachers' adaptation to these new and challenging contexts.

Authors' contributions

María-José Mera-Lemp. Conceptualization; Formal analysis; Investigation; Methodology; Supervision; Writing (original draft); Writing (review & editing).

Javier Torres-Vallejos. Formal analysis; Methodology; Writing (review & editing).

Florencia Guglielmetti-Serrano. Conceptualisation; Investigation; Project administration; Writing (original draft); Writing (review & editing).

References

Abacioglu, C. S., Zee, M., Hanna, F., Soeterik, I. M., Fischer, A. H., & Volman, M. (2019). Practice what you preach: The moderating role of teacher attitudes on the relationship between prejudice reduction and student engagement. *Teaching and Teacher Education*, *86*, 102887. https://doi.org/10.1016/j.tate.2019.102887

Abós, Á., Haerens, L., Sevil, J., Aelterman, N., & García-González, L. (2018). Teachers' motivation in relation to their psychological functioning and interpersonal style: A variable- and person-centered approach. *Teaching and Teacher Education*, *74*, 21-34. https://doi.org/10.1016/j.tate.2018.04.010

Agirdag, O., Van Houtte, M., & Avermaet, P. (2012). Why does the ethnic and socio-economic composition of schools influence math achievement? The role of sense of futility and futility culture. *European Sociological Review*, *28* (3), 366-378. https://doi.org/10.1093/esr/jcq070

Aierdi, X., Álvarez, I., Bonelli, L., Garaizabal, C. & Wagman, D. (2004). *Inmigración. Nuestros miedos e inseguridades [Immigration. Our fears and insecurities]*. Tercera Prensa.

Alexandra, V. (2018). Predicting CQ development in the context of experiential cross-cultural training: The role of social dominance orientation and the propensity to change stereotypes. *Academy of Management Learning & Education*, *17* (1), 62-78. https://doi.org/10.5465/amle.2015.0096

Akcaoğlu, M. Ö., & Kayiş, A. R. (2021). Teacher candidates' multicultural attitude and self-efficacy: the mediating role of intercultural sensitivity. *Cukurova University Faculty of Education Journal*, *50* (2), 1241-1262. https://dergipark.org.tr/en/pub/cuefd/issue/65577/884922

Allport, G. W. (1968). The historical background of modern social psychology. In G. Lindzey, & E. Aronson (Eds.), *Handbook of social psychology* (vol. 1) (pp. 1-80). Addison-Wesley.

Alonso, M. L. (2011). *El impacto de la diversidad sociocultural en la escuela primaria desde la perspectiva del docente: perfil psico-social y (re) construcción de significados* [Doctoral Dissertation, Universidad del País Vasco]. Archivo Digital Docencia Investigación. http://hdl.handle.net/10810/12162

Argüello-Gutiérrez, C., López-Rodríguez, L., & Vázquez, A. (2024). The effect of moral foundations on intergroup relations: The salience of fairness promotes the acceptance of minority groups. *Social Psychological and Personality Science*, *15* (1), 93-105. https://doi.org/10.1177/19485506231162161

Bennet, M. J. (1986). A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations*, *10* (2), 179-196. https://doi.org/10.1016/0147-1767(86)90005-2

Beniscelli, L., Riedemann, A., & Stang, F. (2019). Multicultural y, sin embargo, asimilacionista. Paradojas provocadas por el currículo oculto en una escuela con alto porcentaje de alumnos migrantes [Multicultural, yet assimilationist. Paradoxes caused by the hidden curriculum in a school, with a high percentage of migrant students]. *Calidad en la educación*, (50), 393-423. https://doi.org/10.31619/caledu.n50.522

Berry, J. W., Kalin, R. & Taylor, D. (1977). *Multiculturalism and ethnic attitudes in Canada*. Ministry of Supply and Services.

Blazevic, I., Bulić, M., & Batarelo Kokic, I. (2023, 30 June). *The relation between intercultural sensitivity and multicultural teaching competency of primary school teachers* [Conference Presentation]. International Conference Didactic Challenges IV: Future Studies in Education, Osijek, Croacia. https://bit.ly/blazevic

Cala, V. C., Soriano-Ayala, E., & López-Martínez, M. J. (2018). Actitudes hacia personas refugiadas y ciudadanía europea inclusiva. Análisis para una propuesta educativa intercultural con el profesorado en formación [Attitudes toward refugees and inclusive European citizenship. Analysis for an intercultural educational proposal with teachers in training]. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, *24* (2), 2. https://doi.org/10.7203/relieve.24.2.13320

Cerón, L., Pérez, M., & Poblete, R. (2017). Percepciones docentes en torno a la presencia de niños y niñas migrantes en escuelas de Santiago: retos y desafíos para la inclusión [Teachers' perceptions regarding the presence of migrant boys and girls in schools of Santiago: Challenges for inclusion]. *Revista latinoamericana de educación inclusiva*, *11* (2), 233-246. https://doi.org/10.4067/S0718-73782017000200015

Chen, G.-M., & Starosta, W. J. (2000). The development and validation of the intercultural sensitivity scale. *Human Communication*, *3*, 1-15. https://digitalcommons.uri.edu/com_facpubs/36/

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, *49* (8), 566-578. https://doi.org/10.3102/0013189X20937240

Crozier, G., & Davies, J. (2008). 'The trouble is they don't mix': Self-segregation or enforced exclusion? *Race Ethnicity and Education*, *11* (3), 285-301. https://doi.org/10.1080/13613320802291173

Dozo, N. (2015). *Gender differences in prejudice: A biological and social psychological analysis* [Doctoral Dissertation, University of Queensland]. UQ eSpace. https://doi.org/10.14264/uql.2015.777

Dubbeld, A., de Hoog, N., den Brok, P., & de Laat, M. (2019). Teachers' multicultural attitudes and perceptions of school policy and school climate in relation to burnout. *Intercultural Education*, *30* (6), 1-19. https://doi.org/10.1080/14675986.2018.1538042

Fisher, M. H. (2011). Factors influencing stress, burnout, and retention of secondary teachers. *Current Issues in Education*, *14* (1), 1. https://cie.asu.edu/ojs/index.php/cieatasu/article/view/658

Flanagan-Bórquez, A., Benavides-Cereceda, C., Fuentes-Madariaga, C., Kraemer-Ibaceta, S., & Sepúlveda-Vicencio, M. (2021). Estudio sobre las experiencias de docentes chilenos que trabajan con estudiantes inmigrantes en escuelas públicas de la región de Valparaíso, Chile [Study on the experiences of chilean teachers who work with immigrant students in public schools in the Valparaiso region, Chile]. *Perspectiva Educacional*, *60* (3), 32-56. https://doi.org/10.4151/07189729-vol.60-iss.3-art.1218

Glock, S., Kovacs, C., & Pit-ten, I. (2019). Teachers' attitudes towards ethnic minority students: Effects of schools' cultural diversity. *British Journal of Educational Psychology*, *89* (4), 616-634. https://doi.org/10.1111/bjep.12248

González, J. A., Villarroel, A. B., & Bastías, L. S. (2023). La educación intercultural en Chile. Estrategias pedagógicas en profesores jefes [Intercultural education in Chile. Pedagogical strategies in head teachers]. *Palimpsesto*, *13* (22), 97-116. https://doi.org/10.35588/pa.v13i22.6129

Grütter, J., Meyer, B., Philipp, M., Stegmann, S., & van Dick, R. (2021). Beyond ethnic diversity: The role of teacher care for interethnic relations. *Frontiers in Education*, *5*, 586709. https://doi.org/10.3389/feduc.2020.586709

rep

Gutentag, T., Horenczyk, G., & Tatar, M. (2018). Teachers' approaches toward cultural diversity predict diversity-related burnout and self-efficacy. *Journal of Teacher Education*, *69* (4), 408-419. https://doi.org/10.1177/0022487117714244

Haenni, A., Wolfgramm, C., Müller, M., Heinzmann, S., & Buholzer, A. (2019). Immigrant students and their teachers. Exploring various constellations of acculturation orientations and their impact on school adjustment. *Intercultural Education*, *30* (5), 478-494. https://doi.org/10.1080/14675986.2019.1586214

Horenczyk, G., & Tatar, M. (2002). Teachers' attitudes toward multiculturalism and their perceptions of the school organizational culture. *Teaching and Teacher Education*, *18* (4), 435-445. https://doi.org/10.1016/S0742-051X(02)00008-2

Instituto Nacional de Estadísticas y Servicio Nacional de Migraciones. (2022). *Informe de resultados de la estimación de personas extranjeras residentes en Chile al 31 de diciembre de 2021. Desagregación nacional, regional y principales comunas [Report on the results of the estimate of foreign residents in Chile as of 31 December 2021. National, regional and main communes disaggregation]*. Instituto Nacional de Estadística; Servicio Nacional de Migraciones. Recuperado el 14 de diciembre de 2023 de https://bit.ly/ineysermig

Jiménez, F., & Fardella, C. (2015). Diversidad y rol de la escuela: discursos del profesorado en contextos educativos multiculturales en clave migratoria [Diversity and roles at schools: Teachers' discourse in multicultural educational contexts]. *Revista Mexicana de Investigación Educativa*, *20* (65), 419-441.

Koc, Y., & Anderson, J. R. (2018). Social distance toward Syrian refugees: The role of intergroup anxiety in facilitating positive relations. *Journal of Social Issues*, *74* (4), 790-811. https://doi.org/10.1111/josi.12299

Khukhlaev, O., Novikova, I., & Chernaya, A. (2022). Interpersonal mindfulness, intergroup anxiety, and intercultural communication effectiveness among international students studying in Russia. *Frontiers in Psychology*, *13*, 841361. https://doi.org/10.3389/fpsyg.2022.841361

Khukhlaev, O. E., Pavlova, O. S., Khakimov, E. R., & Khukhlaeva, O. V. (2021). Measuring intercultural competence in teachers: Development and evaluation of situational judgment test. *Psychological Science and Education*, *26* (6), 46-57. https://bit.ly/Khukhlaev

Kim, H., & Connelly, J. (2019). Preservice teachers' multicultural attitudes, intercultural sensitivity, and their multicultural teaching efficacy. *Educational Research Quarterly*, *42* (4), 3-20.

León, B., Mira, A. R., & Gómez, T. (2007). An evaluation of opinions concerning immigration and multiculturality in the school for teacher training. *Electronic Journal of Research in Educational Psychology*, *5* (12), 259-282. https://dspace.uevora.pt/rdpc/handle/10174/10701

Makarova, E., 't Gilde, J., & Birman, D. (2019). Teachers as risk and resource factors in minority students' school adjustment: An integrative review of qualitative research on acculturation. *Intercultural Education*, *30* (5), 448-477. https://doi.org/10.1080/14675986.2019.1586212

Martínez-Zelaya, G. P., Mera-Lemp, M. J., & Bilbao, M. (2020). Preferencias aculturativas y su relación con la sensibilidad intercultural y el bienestar [Acculturative preferences, intercultural sensitivity and well-being: Chileans as a receiving society to Latin American immigration]. *Revista de Psicología*, *29* (2). https://doi.org/10.5354/0719-0581.2020.55961

McCorkle, W., & Rodriguez, S. (2023). Levels of nationalism among middle and high school social studies teachers: Implications for promoting equity for immigrant students and with educators. *Journal of Social Studies Research*, *47* (2), 92-107. https://doi.org/10.1016/j.jssr.2022.01.004

Mendes, W. B., & Koslov, K. (2013). Brittle smiles: Positive biases toward stigmatized and outgroup targets. *Journal of Experimental Psychology: General*, *142* (3), 923-933. https://doi.org/10.1037/a0029663

Mera-Lemp, M. J., Bilbao, M., & Basabe, N. (2020). School satisfaction in immigrant and Chilean students: The role of prejudice and cultural self-efficacy. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.613585

Mera-Lemp, M. J., Bilbao, M., Martínez-Zelaya, G., & Garrido, A. (2021). Los estudiantes de carrera en educación preescolar ante la inmigración latinoamericana en Chile [Early childhood education students and Latin American immigration in Chile]. *Revista Electrónica de Investigación Educativa*, *23*,e23, 1-14. https://doi.org/10.24320/redie.2021.23.e23.3659

Ministerio de Educación, Centro de Estudios. (2023). *Informe estadístico del sistema educacional con análisis de género 2023 (con datos del año 2022). Documento de trabajo 30 [Statistical report on the education system with gender analysis 2023 (with data from 2022). Working document 30].* Ministerio de Educación. https://bibliotecadigital.mineduc.cl/handle/20.500.12365/19641

Monreal, T., & McCorkle, W. (2021). Social studies teachers' attitudes and beliefs about immigration and the formal curriculum in the United States south: A multi-methods study. *The Urban Review*, *53* (1), 1-42. https://doi.org/10.1007/s11256-020-00561-3

Morales, M., Ortiz, S., Cárdenas, K., López, V., & Torres-Vallejos, J. (2022). Caracterización de la gestión de la convivencia escolar en Chile [Characterization of school coexistence in Chile]. *Perspectiva Educacional*, *61* (3), 80-101. https://doi.org/10.4151/07189729-Vol.61-Iss.3-Art.1286

Nguyen, L. A., & Boreczky, Á. (2022). Diversity and Hungarian teachers' views on nationalism, migration and multiculturalism. *Intercultural Education*, *33*, 264-281. https://doi.org/10.1080/14675986.2022.2070376

Obach, A., Sirlopú, D., & Urrutia, C. (2021). Imaginario social de docentes y profesionales de salud de tres colegios de Santiago sobre el cuerpo y la sexualidad de escolares migrantes latinoamericanos [Social imaginary of teachers and health professionals from three schools in Santiago about body and sexuality of Latin American migrant students]. *Revista Chilena de Antropología*, (43), 216-232. https://doi.org/10.5354/0719-1472.2021.64441

Organización Internacional del Trabajo. (2023). *Estrategia regional de la OIT de migración laboral y movilidad humana en América Latina y el Caribe 2023-2030 [ILO regional strategy on labour migration and human mobility in Latin America and the Caribbean 2023-2030].* Organización Internacional del Trabajo. https://bit.ly/oitmigraciones

Ortega, L., Boda, Z., Treviño, E., Arriagada, V., Gelber, D., & Escribano, M. del R. (2020). The centrality of immigrant students within teacher-student interaction networks: A relational approach to educational inclusion. *Teaching and Teacher Education*, *95*, 103126. https://doi.org/10.1016/j.tate.2020.103126

Pastor-Vicedo, J. C., Contreras-Jordán, O. R., Gil-Madrona, P., & Cuevas-Campos, R. (2016). Estereotipos y prejuicios hacia los inmigrantes en los futuros profesores de educación física [Stereotypes and prejudices in pre-service physical education teachers towards immigrants]. *Cuadernos de Psicología del Deporte*, *16* (3), 123-136. https://bit.ly/pastorvicedo

Pavez-Soto, I., Ortiz-López, J., & Domaica-Barrales, A. (2019a). Percepciones de la comunidad educativa sobre estudiantes migrantes en Chile: trato, diferencias e inclusión escolar [Perceptions of the educational community about migrant students in Chile: Treatment, differences and school inclusión]. *Estudios Pedagógicos (Valdivia)*, *45* (3), 163-183. https://doi.org/10.4067/S0718-07052019000300163

Pavez-Soto, I., Ortiz-López, J. E., Sepúlveda, N., Jara, P., & Olguín, C. (2019b). Racialización de la niñez migrante haitiana en escuelas de Chile [Racialization of Haitian migrant childhood in Chilean schools]. *Interciencia*, *44* (7), 414-420. https://bit.ly/pavezsoto

Prino, L., Sclavo, E., & Quaglia, R. (2008). Prejudice in school: A research among primary school teachers [Los prejuicios en la escuela: una investigación entre profesores de primaria]. *European journal of education and psychology*, *1* (3), 27-37. https://bit.ly/prinosclavo

Riedemann, A., Stefoni, C., Stang, F., & Corvalán, J. (2020). Desde una educación intercultural para pueblos indígenas hacia otra pertinente al contexto migratorio actual. Un análisis basado en el caso de Chile [From an intercultural education for indigenous peoples to another one focused on the current migration context. An analysis based on the Chilean case]. *Estudios atacameños*, (64), 337-359.

Salas, N., Castillo, D., San Martín, C., Kong, F., Thayer, L. E., & Huepe, D. (2017). Inmigración en la escuela: caracterización del prejuicio hacia escolares migrantes en Chile [Immigration in school: Characterization of prejudice towards migrants schoolchildren in Chile]. *Universitas psychologica*, *16*, 161-175. https://doi.org/10.11144/javeriana.upsy16-5.iecp

Sánchez-Sánchez, G. I., & Jara-Amigo, X. E. (2018). De la formación inicial al trabajo docente: comprensión de la trayectoria [From initial teacher training to teacher work: Understanding the teacher's career path]. *Revista Educación*, 1-20. https://doi.org/10.15517/revedu.v42i2.23996

Sanhueza, S. V., Friz, M. C., & Quintriqueo, S. (2014). Estudio exploratorio sobre las actitudes y comportamiento del profesorado de Chile en contextos de escolarización de alumnado inmigrante [Exploratory research on the attitudes and behaviors of teachers from Chile in the immigrant student school setting]. *Revista electrónica de investigación educativa*, *16* (3), 148-162. https://bit.ly/sanhueza

Stefoni, C. & Corvalán, J. (2019). Estado del arte sobre inserción de niños y niñas migrantes en el sistema escolar chileno [The state of the art on the insertion of migrant children in the Chilean school system]. *Estudios pedagógicos (Valdivia)*, *45* (3), 201-215. https://doi.org/10.4067/S0718-07052019000300201

Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *Personality and Social Psychology Review*, *18*(3), 239-255. https://doi.org/10.1177/1088868314530518

Stephan, W. G., & Stephan, C. W. (1993). Cognition and affect in stereotyping: Parallel interactive networks. In D. Mackie, & D. Hamilton (Eds.), *Affect, cognition and stereotyping. Interactive processes in group perception* (pp. 111-136). Academic Press.

Stephan, W. G., & Stephan, C. W. (2013). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23-45). Lawrence Erlbaum Associated.

Tatar, M., & Horenczyk, G. (2003). Diversity-related burnout among teachers. *Teaching and Teacher Education*, *19* (4), 397-408. https://doi.org/10.1016/S0742-051X(03)00024-6

Tijoux, M. E., & Zapata-Sepúlveda, P. (2019). Niños y niñas hijos de inmigrantes en Arica y Parinacota, Chile: propuesta metodológica para el estudio de su vida cotidiana en las escuelas. *Interciencia*, *44* (9), 540-548. https://bit.ly/tijouxzapata

Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial bias in perceptions of others' pain. *PLOS ONE*, *7* (11), e48546. https://doi.org/10.1371/journal.pone.0048546

Tyler, K. M., Wade Boykin, A., & Walton, T. R. (2006). Cultural considerations in teachers' perceptions of student classroom behavior and achievement. *Teaching and Teacher Education*, *22* (8), 998-1005. https://doi.org/10.1016/j.tate.2006.04.017

Ulbricht, J., Schachner, M. K., Civitillo, S., & Noack, P. (2022). Teachers' acculturation in culturally diverse schools. How is the perceived diversity climate linked to intercultural self-efficacy? *Frontiers in Psychology*, *13*, 953068. https://doi.org/10.3389/fpsyg.2022.953068

María-José MERA-LEMP, Javier TORRES-VALLEJOS and Florencia GUGLIELMETTI-SERRANO

Ulrey, K. L., & Amason, P. (2001). Intercultural communication between patients and health care providers: An exploration of intercultural communication effectiveness, cultural sensitivity, stress, and anxiety. *Health Communication*, *13* (4), 449-463. https://doi.org/10.1207/s15327027hc1304_06

Valcke, B., Dierckx, K., Desouter, L., Van Dongen, S., Van Hal, G., & Van Hiel, A. (2022). The contribution of teacher, parental and peer support in self-reported school and general well-being among ethnic-cultural minority and majority youth. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1051143

Zanga, G., & De Gioannis, E. (2023). Discrimination in grading: A scoping review of studies on teachers' discrimination in school. *Studies in Educational Evaluation*, *78*, 101284. https://doi.org/10.1016/j.stueduc.2023.101284

Authors' biographies

María-José Mera-Lemp. She has a doctorate in Psychology from the University of the Basque Country / Euskal Herriko Unibertsitatea. Currently, she is an academic and researcher at the Escuela de Ciencias Jurídicas y Sociales de la Universidad Viña del Mar, Chile. Her research lines address intergroup relations between immigrants and members of host societies in educational contexts, with focus on the study of the impact of intercultural competences on teachers, students and parents. She has presented various communications and seminars on these topics. She has also published her works in scientific journals and books of proven quality. She has taught undergraduate and graduate courses at various Chilean universities. She is currently the principal researcher of the study "Immigrant guardians in school communities: understanding the relevance of the climate of cultural diversity, role construction and involvement in school", funded by the National Agency for Research and Development of the Government of Chile.

https://orcid.org/0000-0001-5763-6913

Javier Torres-Vallejos. He holds a PhD in Psychology, Social Transformations and Subjectivity from the Pontifical Catholic University of Valparaíso, and a PhD in Psychology, Health and Quality of Life from the University of Girona, Spain. He is an academic at the School of Psychology at the Universidad Santo Tomás, Chile, and an associate researcher at the CIELO Research Center: Family, Work and Citizenship at the same university. His lines of research focus on the child and youth population, particularly students, coexistence and school violence, well-being, educational trajectories and research methodology with an emphasis on a quantitative approach. He has worked as an undergraduate and graduate teacher at various Chilean universities. He is currently the principal investigator of the study "Where I come from and how I am: Educational trajectories and subjective well-being in high school students", funded by the National Agency for Research and Development of the Government of Chile; and co-investigator of the project "School violence and school coexistence: Analysis of trends over time, association with mental health and effects on educational trajectories", funded by the Ministry of Education of Chile.

https://orcid.org/0000-0002-4229-7768

Florencia Guglielmetti-Serrano. She holds a master's degree in Psychology Research from the Alberto Hurtado University. Her lines of research focus on migration and well-being in Chile. She has published her work in scientific journals on child protection program workers and negative emotions and their effects on political decisions. She has also presented various papers on these topics at different conferences. Currently, she is a professor of Research Methodology and Quantitative Methods at the Gabriela Mistral University and the University of Viña del Mar. In addition, she works as coordinator of the projects "Immigrant guardians in school communities: understanding the relevance of the climate of cultural diversity, role construction and involvement in school", "Relationships between preschool teachers and immigrant guardians: what explains school involvement in early education?", and "Attitudes towards multiculturalism as a facilitator of user care in the municipal space. A look at the encounter between employees and Latin American immigrants in municipal social services"; all financed by the National Agency for Research and Development of the Government of Chile.

https://orcid.org/0009-0006-7682-3603

Parenting dimensions in migrant and Spanish families. Self-determination theory and positive parenting

Dimensiones educativas parentales en familias migrantes y españolas. Teoría de la autodeterminación y parentalidad positiva

María-Elena RIVOIR-GONZÁLEZ. Predoctoral Researcher. Universidad de Oviedo, Spain (*rivoirmaria@uniovi.es*).

Carmen-María FERNÁNDEZ-GARCÍA, PhD. Associate Professor. Universidad de Oviedo, Spain (*fernandezcarmen@uniovi.es*).

Susana TORÍO-LÓPEZ, PhD. Associate Professor. Universidad de Oviedo, Spain (*storio@unioivi.es*).

Abstract:

In keeping with the principles of positive parenting, understood as parental actions based on the best interests of children and aimed at their overall development, self-determination theory (SDT) posits that it is the family that must satisfy three basic psychological needs: autonomy, relatedness, and competence. Owing to migratory movement, the wide range of cultures present in Spain, and the fact that the family is the preferred place to experience human relationships, this context seems to be an ideal setting for testing the apparent universality of SDT. Hence, the main objective of this study was to examine the parenting dimensions: autonomy (autonomy support vs psychological control), relatedness (warmth vs rejection), and competence (structure vs chaos) and their influence on children's behaviors in families from a wide range of international backgrounds. The sample comprised a total of 3428 parents and 1785 children attending primary education (7-13 years old) in eight autonomous communities in Spain. Information was collected using three standardized questionnaires, adapted and validated for the Spanish context. The results show that, regardless of family origin, there were higher scores in positive parenting dimensions (autonomy support, warmth, and structure) than negative dimensions (psychological control, rejection and chaos). The findings also indicate that warmth had a positive effect on

children's behaviors while psychological control had a negative effect, as well as that there were some differences between certain geographical zones in warmth, structure, and chaos. Our study provides confirmation of the universality of the precepts of SDT and positive parenting in families from a wide range of geographical and cultural backgrounds in the Spanish context.

Keywords: family; migration; self-determination theory; positive parenting; family education; parental education.

Resumen:

En concordancia con los principios de la parentalidad positiva (entendida esta como la actuación de los progenitores basada en el interés superior de los menores y dirigida a su desarrollo integral), la teoría de la autodeterminación (TAD) plantea que la familia debe satisfacer tres necesidades psicológicas básicas: autonomía, relación y competencia. Debido a los movimientos migratorios, a la gran heterogeneidad cultural presente en España y al hecho de que la familia se constituye como ámbito preferente para vivenciar las relaciones humanas, dicho contexto parece un escenario idóneo para comprobar la aparente universalidad de la TAD. De ahí que el principal objetivo de este trabajo fuera analizar las dimensiones educativas parentales: autonomía (apoyo a la autonomía *vs.* control psicológico), relación (afecto *vs.* rechazo) y competencia (estructura *vs.* caos), y su influencia en la conducta de los menores en familias de procedencias internacionales diversas. La muestra se compuso de un total de 3428 progenitores y 1785 menores que cursaban educación primaria (7-13 años) en ocho comunidades autónomas españolas. Para recoger la información, se emplearon tres cuestionarios estandarizados adaptados y validados para el contexto español. Los resultados evidencian que, con independencia del origen familiar, las dimensiones educativas positivas (apoyo a la autonomía, afecto y estructura) obtuvieron puntuaciones más elevadas que las negativas (control psicológico, rechazo y caos). También denotan la influencia positiva del afecto y negativa del control psicológico en la conducta de los menores, así como ciertos matices y diferencias entre determinadas zonas geográficas en relación con el afecto, la estructura y el caos. Por tanto, nuestra contribución confirma el cumplimiento y la universalidad de los preceptos establecidos por la TAD y la parentalidad positiva en familias de procedencias geográficas y culturales muy diversas dentro del contexto español.

Palabras clave: familia, migración, teoría de la autodeterminación, parentalidad positiva, educación familiar, educación parental.

1. Introduction

According to the classical concept of socialization from Rocher (1973), human beings need to learn social and cultural norms to adapt to their surroundings, incorporating them into their personalities through the influence of various social agents. In this regard, families are people's primary in-group and socializing agent, and as such are vital to their learning, survival, and overall development, as well as being one of the main

sources of social cohesion, integration, and development. This socializing role of the family is particularly important during childhood through primary socialization. This is a time when children learn their parents sociocultural and parenting rules through observation, imitation, and reward, while at the same time, the parents (in their secondary socialization processes) learn to play their respective roles (mother/father) throughout their lives, as described by Schaffer's (1989) mutuality model. During these socialization processes, motivated learning, the assimilation of social and cultural norms through behavioral reinforcement, will play a very important role. Llopis and Llopis (2003) are adamant about this idea when referring to parenting, a process through which parents shape the behaviors that they feel are proper and desirable from their children.

It is important to emphasize that in this study we focus on parents and children as basic members of the nuclear or conjugal family as referred to by Durkheim (1892). He attributed an organic kind of solidarity to the relationships in modern societies, motivated by greater division of labor, smaller families, and greater independence from the extended family in favor of other individuals and institutions. This is a key consideration in contexts such as Spain and is related to a growing trend towards individualism or the privatization of morals. Added to that, there is a greater awareness of democratic principles such as equality and tolerance, which means that families no longer unfailingly follow the tradition-

al rules. This has led to a fall in the number of marriages and a rise in cohabitation (Cea-D'Ancona, 2007).

The changing profile of Spanish families confirmed by recent studies (Instituto de Política Familiar, 2021; OECD, 2022) is characterized by low fertility rates, an aging society, the dominance of double source income families, greater involvement of parents in household chores and time dedicated to children, and an intergenerational support. In many dimensions, people living in Spain benefit from similar or better well-being levels than OECD countries averages. However, a significant part of the population struggles with adversities such as high poverty rates, a weak labor market, the cost of housing and significant difficulties in reconciling family and work life. Therefore, the design of policies and measures such as increasing economic benefits and employment quality, promoting paternity/maternity leave, guaranteeing equal opportunities, reducing early school leaving, ensuring the right of housing (prioritizing families with children) or improving children well-being can enable a positive impact for all families (Llano, 2023; Jorquera & Del Moral, 2020).

1.1. Migration and families
When it comes to migrant families, the role of parenting becomes even more important, if possible, given that families make up a large proportion of the migrant population (OECD, 2017). Added to that are the more limited social and emotional networks in the destination country,

due in part to how far away the extended family is, and to the idea that the family environment is the preferred place for experiencing stronger human relationships (Martín, 2008, as cited in Fernández-Hawrylak et al., 2016).

We must bear in mind that most of the migrant population (68%) in Spain originate from non-EU countries. The most common countries of origin are in Latin America (30%) and Africa (24%) (Instituto Nacional de Estadística [INE], 2023a). The most common nationalities are Moroccan (17%), Romanian (12%), and Colombian (6%). In the migrant population, the most common age range is 35-29 years old, compared to the Spanish population, where the most common is 45-49 years old (INE, 2023b). This is one reason why the migrant population is a particularly important factor in combatting the aging of the local population. It is also a relevant fact since the most common age group in the migrant population concurs with important work activity and with motherhood and paternity, central aspects in the objectives of the present study.

This multicultural social and educational reality involves what Donati (2003), in terms of family socialization, called a re-evaluation of the family, in so much as migrant families, as we have seen, largely come from pre-modern societies or societies that have followed other routes of modernization in which the family is generally considered to be a strong social structure that exhibits notable internal solidarity. This is reflect-ed in the importance of social capital in terms of connections (family and friendship networks) in the destination country that facilitate settlement and social integration (Páez-de la Torre, 2020). Unlike native-born families, characterized by more fragmented lives (Donati, 2003) (an organic type of solidarity as noted by Durkheim in 1892) and a weakening of the family institution, families that come from more traditional societies have very strong family ties (Pérez-Díaz et al., 2001), which fit the characteristics of the mechanics of solidarity posited by Durkheim (1892) and reflect a strong sense of collectivism. Values such as kinship, sharing with others, helping members of the family in the broadest sense (the extended family), and respect for elders (grandparents) as sources of authority are often felt to be at the core of African, Latin American, Chinese, and Indian households, as noted by Habiyakare (2003), Polanco (2003), Guo (2003), and Chittayath (2003). In Africa, China, and India in particular, we can talk about principally community-based societies that, with the advance of urbanization and capitalism, have been pushed back by the individualist societies that are more widespread nowadays in the west.

1.2. Self-determination theory as a conceptual reference framework

Within the analysis of parenting dimensions, self-determination theory (SDT) is a powerful explanatory concept that comes together with the principles of positive parenting, which is parental action based on the greater interest of

the child and aimed at developing their abilities, without violence, through recognition and setting limits (Consejo de Europa, 2006; Torío-López et al., 2022). From this perspective, it is the family that should satisfy the three basic psychological needs referred to by SDT (Deci & Ryan, 1985): autonomy, relatedness, and competence.

In terms of the first of those three, *autonomy* support is the ideal parenting dimension, encouraging children to take initiative and express their opinions, and giving them space to solve their own problems (Grolnick & Pomerantz, 2009). In contrast, psychological control through reward or punishment, as defined by Barber (1996, as cited in Soenens & Vansteenkiste, 2010), is an intrusion into children's thoughts via manipulative parenting techniques such as inducing guilt or shame and withdrawal of love, which harm children's autonomy. Baumrind (1966) identified three types of control: authoritarian, authoritative, and permissive. Authoritative is the most desirable in terms of positive parenting. Studies in different cultural samples have shown that the psychologically controlling parenting profile is frequently associated with negative results in children (Soenens & Vansteenkiste, 2010; Yu et al., 2014).

The second need, *relatedness*, refers to a feeling of belonging and is facilitated by communicating respect and affection (Deci & Ryan, 1985). It has been widely demonstrated that when adults offer attention, warmth, and emotional support,

children's development sees notable benefits. In addition, parental lack of affection has a negative effect on children's physical and intellectual development and may lead to potential emotional disorders that make subsequent social relationships more difficult, as well as affecting the development of their adult personality (Abad, 1993; Bowlby, 1951; Mestre et al., 1997; Torío-López et al., 2008).

Lastly, *competence* is about the need to acquire certain tools that will allow children to do certain tasks and overcome adversity (Torío-López et al., 2022). Ryan and Deci (2020) refer to the sensation of being able to succeed and grow and indicated that this is better satisfied in well-structured (non-chaotic) environments with positive feedback and opportunities for growth. Farkas and Grolnick (2010) note six components for parents or guardians that characterized structure in the home: clear rules, guidelines and expectations; predictable consequences; task-focused feedback; opportunity to meet expectations; rationales for rules and expectations; and authority.

Moreover, the satisfaction and frustration of these basic psychological needs would be related to adolescents' well-being and psychological maladjustment (Van der Kaap-Deeder et al., 2017; Rodríguez-Meirinhos et al., 2020). Without any doubts, prosocial disposition can be considered as a protective factor or modulator of aggressive behavior and emotional instability. Tur-Porcar et al.

(2018) verify the positive relationship of prosocial behavior with attachment (father and mother), functional coping and peers acceptance. Likewise, they showed the negative relationship of prosocial behavior with abandonment (father and mother), emotional instability, aggressiveness, dysfunctional coping and peers rejection.

SDT suggests that these three psychological needs are universal and part of the human condition, even in more collectivist cultures, as Soenens et al. (2015) noted. According to Donati (2003), the process of intra-occidental and intra-continental cultural dialog that we engage in nowadays, which includes ingredients such as identity and family relationships, is a huge opportunity for any comparison between family cultures to also show what they have in common, regardless of origin. In this regard, Spain, with the wide variety of cultures that families here have originated from, is an ideal setting in which to test this apparent universality. Given that, the following objectives for the study were established: a) analyze the parenting dimensions of migrant families in Spain using SDT as a theoretical framework: autonomy (autonomy support vs psychological control), relatedness (warmth vs rejection), and competence (structure vs chaos); b) examine the possible differences and similarities in the parenting dimensions in terms of family origin; c) determine children's opinions about the autonomy provided by their parents and whether this corresponds to the parents' opinions in the migrant families; and d)

predict the children's behaviors according to the parenting dimensions, considering family origin and educational attainment.

We hypothesize that the main precepts of SDT are universal and also apply in the Spanish context, regardless of family origin (Spanish/migrants), and that origin will not determine differences in children's behaviors. This information should help us to clarify guidelines, and lead to more accurate diagnoses from which interventions can be designed that are aimed at the most common family profiles in Spain.

2. Method

The methodology used in the study followed a non-experimental, *ex post facto* design.

2.1. Participants

The sample was made up of 5213 participants. 85.8% of the families were Spanish and 14.2% migrants. The current migrant population in Spain is around 12% (INE, 2023c), which is a similar proportion in our study, indicating that this is a representative sample. The mean age of the Spanish parents was 43 years old; the mean age of the migrant parents was 42.9. The children were aged between 7 and 13 years old (in the 3rd to the 6th year of primary education). The families were resident in eight of the autonomous communities in Spain, and the migrant families (like the general migrant population in Spain) came from a wide range of geographical

regions: North Africa (22.1%); Venezuela, Colombia, and Brazil (18.7%); the Andean States (12.8%); Eastern Europe (11.4%); Western Europe (10.7%); the Southern Cone (10.4%); Central America and Mexico (7.6%); Sub-Saharan Africa (2.4%); Asia (2.1%); and North America and Australia (1.7%). 847 (47.5%) of the sample were boys and 932 (52.3%) girls. The sample was formed by 1627 (48.3%) fathers with an average age of 45.2 years and by 1742 (51.7%) mothers with an average age of 42.73 years. A total number of 1738 families formed the sample, being most of them nuclear or conjugal and heterosexual. Both parents participated in the study.

We found a relationship between the type of school the children attended and the family type (Fisher's test, p <.001). Although most of the children attended state-funded public schools (78.4% of Spanish families [SF], 69.1% of migrant families [MF]), a higher proportion of children from migrant families than Spanish families went to *concertado* schools [these are semi-private schools that receive some state funding but are administratively independent] (20.5% SF, 30.2% MF), whereas a higher proportion of children from Spanish families attended private schools (1.1% SF, 0.7% MF), which were the minority. In terms of parental educational attainment, there was also a relationship with family type (Pearson's chi-squared test, p <.001). In the Spanish families, 45% had secondary-school qualifications, 42.7% had university-level qualifications, and 12.2% had only a primary education. In

the migrant families the picture was different: 54.6% had a secondary-school education, a higher proportion (18.9%) had only a primary education, and a smaller proportion (26.5%) had a university education. These data are in line with Reher et al. (2007), who note a moderate level of education in the migrant population, and with Iglesias et al. (2020), who indicate that migrants' cultural capital was significantly lower than the native population's. According to gender differences and level of studies, it can be observed that an important number of fathers and mothers had secondary-school qualifications 47.4% fathers and 45.7% mothers) and university-level qualifications (35.4% fathers and 44.2% mothers). 17.2% of fathers and 10.1% mothers had only a primary education. It should be noted that the percentage of mothers with university-level qualifications was almost 10 points higher than fathers' one.

Looking at the migrant families' incomes, lower incomes were more common: <15 000 euros (44.4%), 15 000-25 000 euros (22%), 25 000-35 000 euros (11.8%). In the Spanish families, the incomes were higher: 15 000-25 000 euros (22.9%), 25 000-35 000 euros (19.8%), >55 000 euros (15.2%) (p <.001).

The majority of parents were in full-time work (75.9% SF, 53.6% MF), whereas it was more common for migrant families to homemakers (14.1%) or unemployed (12.1%) than for Spanish parents (5.4% and 6.1%, respectively) (p <.001). This sociodemographic picture is in line with

reports from the Spanish Economic and Social Council (CES España, 2019): wage gaps between foreigners and Spanish workers, and a worse position in the job market for the migrant population, with higher rates of unemployment. The European Anti-Poverty Network in Spain (EAPN España, 2021) noted that 36.3% of EU non-nationals living in Spain and 49.5% of those from the rest of the world were living below the poverty line. This is consistent with the most commonly indicated incomes from the migrant families in our study: <15 000 euros (44.4%) and 15 000-25 000 euros (22%), close to the poverty line. The Spanish charities Foessa Foundation and Cáritas (2022) have also noted this pattern, which has only been made worse by the pandemic, highlighting nationality and ethnicity as key factors in social exclusion. The various sources mentioned underscore the representativeness of our sample in Spain, both quantitatively and qualitatively, with regard to the sociodemographic characteristics described above.

2.2. Instruments

The data were collected via standardized questionnaires that had been adapted and validated for the Spanish context. These were completed by the parents and children from the Spanish and migrant families:

- Parents as a Social Context Questionnaire (PSCQ) (Skinner et al., 2005; Spanish adaptation by Inda-Caro et al., 2023). This is a questionnaire for parents to complete individually, comprising 31 items that evaluate

the parenting model based on SDT with three dimensions (Skinner et al., 2005): autonomy support (6 items; e.g., "I encourage my child to express their opinions, even when they differ from mine") vs psychological control (8 items; e.g., "I do not allow my child to decide too many things for themself"); warmth (6 items; e.g., "I do special things with my child") vs rejection (5 items; e.g., "My child fights me at every turn"); structure (5 items; e.g., "I make it clear to my child what will happen if they do not follow the rules") vs chaos (6 items; e.g., "My child needs more time than I have available"). The items are presented with a four-point Likert scale (from 1 = completely disagree to 4 = completely agree). Cronbach's alpha was .916. The Bartlett test indicated that the principle of sphericity was fulfilled (χ^2 = 17,543.8, p = .000) and the Kaiser-Meyer-Olkin test obtained a value of .883 (Inda-Caro et al., 2023).

- Prosocial Behavior Scale (Caprara & Pastorelli, 1993; Spanish adaptation by Tur, 2003). This instrument is completed by children and comprises three scales:

  ◦ The first one measures the child's altruistic behavior through 15 items with three responses: rarely, sometimes, never (e.g., "I try to help others"). Cronbach's alpha was .72 for the Spanish version (Tur-Porcar et al., 2018);

  ◦ The second one is the Emotional Instability Scale (Caprara &

Pastorelli, 1993; Spanish adaptation by Del Barrio et al., 2001), which has 20 items evaluating children's behavior related to lack of control, impulsivity, and emotionality in social contexts. The items have three response options indicating how often each behavior occurs: rarely, sometimes, never (e.g., "I'm impatient"). Alpha for the Del Barrio et al. (2001) version, with children aged 7-10 years old, was .74 (2001); for the Tur-Porcar et al. (2018) adaptation, with children aged 7-12, it was .81.

◦ The third one is the Physical and Verbal Aggressiveness Questionnaire (Caprara & Pastorelli, 1993; Spanish adaptation by Del Barrio et al., 2001). This has 20 items evaluating behavior aimed at physically or verbally hurting others, each with three response options: *rarely*, *sometimes*, *never* (e.g., "I hurt my classmates"). Alpha for the Del Barrio et al. (2001) version, with children aged 7-10 years old, was .84 (2001); for the Tur-Porcar et al. (2018) adaptation, with children aged 7-12, it was .89.

• Perceived Parental Autonomy Support Scale (PPASS) (Mageau et al., 2015; Spanish adaptation by Inda-Caro et al., 2022). This scale is completed by children, and consists of 18 items, each with four response options: *almost never*, *not very often*, *quite often*, and *almost always*. It measures the

children's perceptions of their parents' support for autonomy (9 items; e.g., "They can put themself in my shoes and understand how I feel") or psychological control (9 items; e.g., "When they want me to stop doing something, they say or do things that make me feel bad"). The original version was translated into Spanish. Cronbach's alpha was .92. The Bartlett test indicated that the principle of sphericity was fulfilled (χ^2 = 2069.0, p <.000) and the Kaiser-Meyer-Olkin test obtained a value of .92 (Inda-Caro et al., 2022).

2.3. Procedure

The instruments were adapted from English to Spanish following the rules of the International Test Commission. Two translations were made from English by members of the research team which were assessed by two experts (T1). Subsequently, they were back-translated into English by two native-speaking translators (T2) to confirm a definitive Spanish version following agreement between the participating members of the research team. Following this adaptation to Spanish, the members of the research team, with the help of area coordinators (teachers from various universities), contacted schools in eight different autonomous communities in Spain to ask them to participate in the study. Those who agreed were sent a letter explaining the study which also included a form for informed consent and authorization from parents. The children completed their questionnaires during normal class time, taking between 25 and 45 minutes

to do so depending on age. Owing to the covid-19 situation, certain modifications were required by the authorities (e.g., online assistance for the teachers in the schools, instruments being administered by form tutors following written instructions from the research group, etc.). Parents completed their questionnaires outside of school and returned the completed instruments to the school in sealed envelopes delivered by their children. None of the participants received any reward or remuneration for their participation in the study, which was approved by the Ethics Committee of Principality of Asturias (Cod. CEIm PAst: no 200/19).

2.4. Data analysis

A descriptive analysis was performed first. The study of the relationships between two variables, performed on the sociodemographic variables to better understand the representativeness of the sample (e.g., educational attainment and family), were examined using Pearson chi-squared and Fisher's test given the failure to verify the hypothesis about expected frequencies. To analyze the parental dimensions and their relationship with geographical areas of origin, we looked at the differences of variables between more than two groups with the ANOVA test or the Kruskal-Wallis test depending on whether normality and/or homoskedasticity was confirmed, along with Dunn's test to determine which groups were significantly different. In order to make the statistical analysis possible and to reduce bias, we chose to group countries of origin based on cul-

turally close geographical regions, namely: Sub-Saharan Africa; North America and Australia; Central America and Mexico; Asia; the Southern Cone; the Andean States; Eastern Europe; Western Europe; North Africa; Venezuela, Colombia, and Brazil.

This gave a balance between achieving a minimum number of data points per category and avoiding overgeneralization. Where there were two groups to compare, we used the Student t test or Welch's test for independent samples (e.g., the relationship between PSCQ autonomy support and family). Following that, we specified ANCOVA models to predict children's behaviors based on the parenting dimensions of autonomy support, psychological control, and warmth. Parents' educational attainment and the family origin were also considered in examining the differences. The level of significance was set at .05. Statistical analysis was done using R software (R Core Team, 2022) version 4.1.3.

3. Results

3.1. Descriptive and correlational analysis

The results from the PSCQ questionnaire given to Spanish parents indicated a high level of autonomy support ($M = 3.74$; $SD = 0.36$) on the 4-point scale; a moderate level of psychological control ($M = 1.92$; $SD = 0.61$); a high level of warmth ($M = 3.41$; $SD = 0.46$) (Table 1), with frequencies as follows: high warmth (50.2%) and low warmth (49.8%) (Table 2); and a moderate level of rejection ($M = 1.74$;

SD = 0.57). There was also a high score in structure (M = 3.43; SD = 0.46) and a moderate score in chaos (M = 1.73; SD = 0.54) (Table 1).

Looking at the scores from the parents in migrant families, there was also a high level of autonomy support, similar to the Spanish parents, with a mean score of 3.77 (SD = 0.32). The mean score for psychological control was also moderate, although slightly higher than the score from the Spanish parents (M = 2.04; SD = 0.64). We found a high score in warmth (M=3.50; SD=0.47) (Table 1), with the following frequency distribu-

tion: high scores (60.4%) and low scores (39.6%) (Table 2); and a moderate level of rejection (M=1.80; SD=0.63). Taking into consideration the great academic consensus regarding the importance of warmth in child development these data were particularly highlighted in the current study. Relative percentages were also provided as complementary data. In both measures these were higher than the Spanish families. In the third dimension, competence, we again saw higher scores than the Spanish parents gave, with a high level in structure (M = 3.53; SD = 0.42) and a moderate level in chaos (M = 1.91; SD = 0.62) (Table 1).

TABLE 1. Mean scores from parents in the parenting dimensions.

PSCQ parenting dimensions	Parents in migrant families			Parents in Spanish families		
	n	Mean	SD	n	Mean	SD
Autonomy support	374	3.77	0.32	2286	3.74	0.36
Psychological control	362	2.04	0.64	2241	1.92	0.61
Warmth	365	3.50	0.47	2290	3.41	0.46
Rejection	367	1.80	0.63	2269	1.74	0.57
Structure	371	3.53	0.42	2279	3.43	0.46
Chaos	361	1.91	0.62	2237	1.73	0.54

TABLE 2. Distribution of frequencies and percentages in warmth.

Level of warmth	Migrant families		Spanish families	
	Freq.	%	Freq.	%
High warmth	267	60.4	1336	50.2
Low warmth	175	39.6	1325	49.8
Total	442	100.0	2661	100.0

Subsequently, we examined the relationships between the parenting dimensions and the different areas of origin of the families. Due to the heterogeneity of the sample, the study focused on geographical areas (Table 3) instead of family origin (Spanish/migrants) (Table 4). We found no statistically significant differences in the relationships between PSCQ autonomy support and geographical area (Kruskal-Wallis test, p = .385), or between PSCQ psychological control and geographical area (Kruskal-Wallis test, p = .134), or between PSCQ rejection and geographical area (Kruskal-Wallis test, p = .108), but we did find significant differences in the relationship between PSCQ warmth and geographical area (Kruskal-Wallis test, p = .003). Dunn's test indicated that certain pairs did produce significant differences. In order of significance, they were: Central America-Mexico and Spain (p = .002), Central America-Mexico and Western Europe (p = .009), North Africa and Spain (p = .029), and Venezuela-Colombia-Brazil and Spain (p = .041). Parents from Central America-Mexico, North Africa, and Venezuela-Colombia-Brazil demonstrated significantly higher levels of warmth than Spanish parents (or parents from Western Europe in the case of Central America-Mexico).

With regard to the final dimension, competence, analysis of the relationship between PSCQ structure and geographical area, we did find significant differences (Kruskal-Wallis test, p <.001). Dunn's test again indicated that there were significant differences between certain pairs. Ranked by significance, they were the Andean States and Spain (p <.001), the Southern Cone and Spain (p = .005), the Andean States and Western Europe (p = .011), Venezuela-Colombia-Brazil and Spain (p = .021), and the Southern Cone and Western Europe (p = .049). Parents from the Andean States, the Southern Cone, and Venezuela-Colombia-Brazil had significantly higher levels of structure than Spanish parents, or than parents from Western Europe in the case of the Andean States and the Southern Cone.

Finally, analysis of the relationship between PSCQ chaos and geographical area, we did find significant differences (Kruskal-Wallis test, p <.001). Dunn's test indicated that there were statistically significant differences between many pairs. In order of significance, namely: Eastern Europe and Spain (p <.001), North Africa and Spain (p = .001), the Andean States and Spain (p = .002), Asia and Spain (p = .002), Asia and Venezuela-Colombia-Brazil (p = .004), Asia and the Southern Cone (p = .006), Asia and Western Europe (p = .008), Asia and North Africa (p = .01), Central America-Mexico and Asia (p = .014), the Southern Cone and Eastern Europe (p = .019), Venezuela-Colombia-Brazil and Eastern Europe (p = .021), Asia and the Andean States (p = .025), and Eastern Europe and Western Europe (p = .034). Parents from North Africa and the Andean States had significantly higher levels of chaos than Spanish parents. Parents from Asia had significantly higher levels than most other regions:

Spain, Venezuela-Colombia-Brazil, the Southern Cone, Western Europe, North Africa, Central America-Mexico, and the Andean States. Parents from East-ern Europe also had higher levels than parents from Spain, the Southern Cone, Venezuela-Colombia-Brazil, or Western Europe (Table 3).

TABLE 3. Relationships between PSCQ parenting dimensions and geographical area.

PSCQ parenting dimension and geographical area	Kruskal-Wallis test	Pairs with significant differences according to Dunn's test
PSCQ autonomy support and geographical area	.385	
PSCQ psychological control and geographical area	.134	
PSCQ warmth and geographical area	<.001	Central America-Mexico and Spain (p =.002) Central America-Mexico and Western Europe (p =.009) North Africa and Spain (p = .029) Venezuela-Colombia-Brazil and Spain (p =.041)
PSCQ rejection and geographical area	.108	
PSCQ structure and geographical area	<.001	Andean States and Spain (p <.001) Southern Cone and Spain (p = .005) Andean States and Western Euroope (p = .011) Venezuela-Colombia-Brazil and Spain (p = .021) Southern Cone and Western Europe (p = .049)
PSCQ chaos and geographical area	<.001	Eastern Europe and Spain (p <.001) North Africa and Spain (p =. 001) Andean States and Spain (p = .002) Asia and Spain (p = .002) Asia and Venezuela-Colombia-Brazil (p = .004) Asia and Southern Cone (p = .006) Asia and Western Europe (p = .008) Asia and North Africa (p = .01) Central America-Mexico and Asia (p = .014) Southern Cone and Eastern Europe (p = .019) Venezuela-Colombia-Brazil and Eastern Europe (p = .021) Asia and Andean States (p = .025) Eastern Europe and Western Europe (p = .034)

We found similar results when we examined whether the parenting dimensions behaved differently for the different levels of the family variable (migrant/Spanish) (Table 4). Looking at the relationship between PSCQ autonomy support and family or PSCQ rejection and family, there were no significant differences between the groups (Welch's test, p = .06) (Welch's test, p = .068). Analyzing the relationship between PSCQ psychological control

and family we did find significant differences (Student t test, p < .001). There was a slightly higher level of parental control in migrant families. A similar result was found for the relationship between PSCQ warmth and family (Student t test, p = .001), with higher levels of warmth in migrant families. The difference was more notable in structure (Welch's test, p <.001), and particularly in chaos (Welch's test, p <.001), with higher scores from the migrant families.

TABLE 4. Relationships between PSCQ parenting dimensions and family.

PSCQ parenting dimension and family	Welch's test / Student t test
PSCQ autonomy support and family	.06
PSCQ psychological control and family	<.001
PSCQ warmth and family	.001
PSCQ rejection and family	.068
PSCQ structure and family	<.001
PSCQ chaos and family	<.001

When looking at the first parenting dimension, autonomy, we also asked the children about their perceptions of their parents' autonomy support and psychological control, and examined whether the children's and parents' views agreed. We found through the correlations (Spearman test) that the association between PPASS autonomy support (children) and PSCQ autonomy support (parents) was on the limit of significance (p = .05). The children's

mean score for autonomy support was moderately high (M = 3.10; SD = 0.58), and lower than the parent's mean score, which was high (M = 3.77; SD = 0.32). There was a significant positive relationship between the two in terms of psychological control (p <.001), with a high level of agreement between the mean scores from PPAS psychological control (children) (M = 1.81; SD = 0.55) and PSCQ psychological control (parents) (M = 1.92; SD = 0.61), both

rep

indicating a moderate level of psychological control.

3.2. ANCOVA models

We created ANCOVA models for children's behaviors (prosocial behavior, emotional instability, and physical or verbal aggressiveness). The predictor variables included in each model were the results from the parenting dimensions autonomy and relatedness, along with two variables for examining differences: parental education and family origin.

As Table 5 shows, we confirmed that children's behavior is not associated with the parenting dimensions and there were no significant differences according to parental education or family origin.

In the second child-related category, we found that high scores for parental warmth were associated with low scores for children's emotional instability ($p = .002$), which were significantly lower in migrant families ($p = .019$) or when the parents had higher educational attainment ($p = .001$). In contrast, high scores in psychological control were associated with high scores in emotional instability ($p < .001$) (Table 6).

With regard to children's verbal or physical aggressiveness, we found that high scores in parental psychological control were associated with higher scores in aggressive attitudes ($p < .001$). There were differences according to parental educational attainment, with parents who had lower levels of education associated with higher scores in physical aggressiveness ($p = .034$), which were significantly higher in-migrant families ($p = .048$) (Table 7).

TABLE 5. ANCOVA coefficients for prosocial behavior with the results of the parenting dimensions as predictors.

Prosocial behavior	Univariate coefficient	Multivariate coefficient
Warmth	0.02 ($p = .053$)	0.01 ($p = .473$)
Rejection	-0,02 ($p = .015$)	-0,01 ($p = .464$)
Psychological control	-0.03 ($p = .001$)	-0.02 ($p = .068$)
Autonomy support	0.03 ($p = .073$)	0.00 ($p = .812$)
Educational attainment		
Primary		
Secondary	-0.00 ($p = .962$)	-0.00 ($p = .858$)
University	0.03 ($p = .056$)	0.03 ($p = .100$)
Family		
Spanish		
Migrant	-0.03 ($p = .044$)	-0.02 ($p = .383$)

TABLE 6. ANCOVA coefficients for emotional instability with the results from parenting dimensions as predictors.

Emotional instability	Univariate coefficient	Multivariate coefficient
Warmth	-0.10 (p <.001)	-0.06 (p =.002)
Rejection	0.07 (p <.001)	0.01 (p =.516)
Psychological control	0.10 (p <.001)	0.08 (p <.001)
Autonomy support	-0.08 (p <.001)	0.00 (p =.958)
Educational attainment		
Primary		
Secondary	-0.02 (p =.297)	-0.01 (p =.539)
University	-0.09 (p <.001)	-0.08 (p =.001)
Family		
Spanish		
Migrant	-0.04 (p =.039)	-0.05 (p =.019)

TABLE 7. ANCOVA coefficients for physical or verbal aggressiveness with the results from parenting dimensions as predictors

Physical or verbal aggressiveness	Univariate coefficient	Multivariate coefficient
Warmth	-0.06 (p <.001)	-0.02 (p =.236)
Rejection	0.04 (p <.001)	0.00 (p =.758)
Psychological control	0.06 (p <.001)	0.05 (p <.001)
Autonomy support	-0.07 (p <.001)	-0.02 (p =.257)
Educational attainment		
Primary		
Secondary	-0.01 (p =.540)	-0.02 (p =.430)
University	-0.04 (p =.046)	-0.04 (p =.034)
Family		
Spanish		
Migrant	-0.02 (p =.191)	-0.04 (p =.048)

4. Discussion and conclusions

In the Spanish context, migrant families demonstrated what might be desirable according to the precepts of SDT put forward by Deci and Ryan (1985) about the main psychological needs and their corresponding parenting dimensions. In the first place, with regard to autonomy, there was a high level of autonomy support and a moderate level of psychological control. Perceived autonomy support is slightly lower in children than in parents but perceived psychological control is similar in parents and children. In addition, there were high levels of warmth, moderate levels of rejection and higher levels of structure than chaos (which was reported at a moderate level).

Making comparisons based on where the families originated, overall, we found similarities, as reported by Soenens et al. (2015). The first similarity was in the levels reported for the various parenting dimensions, which had comparable scores regardless of family origin, especially in autonomy support. Furthermore, there were higher scores in the dimensions that are considered positive for children's development (autonomy support, warmth, and structure) than in those considered negative (psychological control, rejection, and chaos). It is worth noting that the scores for psychological control, rejection and chaos were not low, but moderate in both Spanish and migrant families. Another similarity was the negative influence of parental psychological control on children's behavior, which was associated with more physical or verbal aggressiveness and greater emotional instability, consist-ent with the findings from Soenens and Vansteenkiste (2010), and Yu et al. (2014). Parental educational attainment also influenced the latter two behaviors. We found that the higher the parents' educational attainment, the lower the children's emotional instability and aggressiveness. A possible cause or relevant factor might be having university-level studies, as a key point which might provide families with more strategies in order to create a positive, non-violent climate. The final similarity was related to the clear positive influence that warmth had on children (Abad, 1993; Bowlby, 1951; Eltink et al., 2024; Mejia-Flores et al., 2024; Mestre et al., 1997; Torío-López et al., 2008). We found that the greater the parental warmth, the lower the children's emotional instability.

In terms of the differences we identified, it is worth noting that although all of the scores in the parenting dimensions were close, as noted above, they were higher in the migrant families than in the Spanish families. The biggest differences were obtained in competence, observing higher levels of chaos in migrant families. In this sense, social and economic aspects need to be considered: the absence of family or social support network, greater job segregation, precarious occupations which often involve strenuous work (CES, 2019; Iglesias et al., 2020). All these circumstances might imply more chaotic environments.

Migrant parents were also slightly more controlling, which is consistent with reports from Henao-Agudelo et al. (2016) and Yu et al. (2014), which influences

higher emotional instability and aggressive attitudes in children. They are also more affectionate, which influences lower emotional instability in children in migrant homes. This pattern is in line with findings from Donati (2003), Pérez-Díaz et al. (2001), Habiyakare (2003), Polanco (2003), Guo (2003), and Chittayath (2003), who suggested that in families from more traditional societies (or societies at different points on the path to modernization) family and group ties are particularly close and important. Our results in terms of specific geographical area of origin were along similar lines. We found differences in warmth, structure, and chaos. Parents from Eastern Europe scored lower in warmth than those from Central America and Mexico, and lower in structure than those from the Southern Cone; parents from North Africa had higher scores in warmth and chaos than Spanish parents; Venezuelan, Colombian, and Brazilian parents scored higher in warmth and structure than Spanish parents; and Asian and Eastern European parents stood out for higher scores in chaos than most other regions. Once again, socioeconomic and labour factors need to be considered to understand chaotic environments. That is, those origins which showed the highest levels in chaos (North Africa, Asia and Eastern Europe) are the ones which experience more vulnerability.

To sum up, and by way of conclusion, we can say that in general terms the precepts laid out in SDT and positive parenting held and we have confirmed their universality in families from a wide range of geographical and cultural origins. These families demonstrated high levels in all of the parenting dimensions that effectively have a positive influence on children's overall development (autonomy support, warmth, and structure), higher than the moderate levels they exhibited in the dimensions that have negative effects (psychological control, rejection, and chaos).

We can confirm our initial hypothesis, the precepts of SDT hold in families regardless of origin, and this origin is not determinant of the differences in children's behaviors. These results, together with the importance of the family as socializing agents, show the need for continued socio-educational intervention, both for families and for the professionals that work with them, to provide resources and tools that encourage and empower positive parenting.

The similarities that we found, which are important, are not incompatible with the differences and details we observed, which are indicative of a notable heterogeneity in terms of the cultural and social diversity the families originated from. In this regard, and one limitation of the study, it is worth noting we must be cautious when generalizing the results, as they are only from the Spanish context, where all of the participants in the study resided.

Authors' contributions

María Elena Rivoir-González: Conceptualisation; Data curation; Methodology; Writing (original draft); Writing (review & editing); Visualisation.

References

Abad, L.V. (1993). Individuo y sociedad: la construcción de la identidad personal [Individual and society: the construction of personal identity]. In M. A. García, G. de la Fuente, & F. Ortega (Eds.), *Sociología de la educación [Sociology of education]* (pp. 23-43). Barcanova.

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development, 37* (4), 887-907. https://doi.org/10.2307/1126611

Bowlby, J. (1951). *Maternal care and mental health: A report prepared on behalf of the World Health Organization as a contribution to the United Nations programme for the welfare of homeless children.* World Health Organization. https://iris.who.int/handle/10665/40724

Caprara, G. V., & Pastorelli, C. (1993). Early emotional instability, prosocial behavior, and aggression: some methodological aspects. *European Journal of Personality, 7*, 19-36. https://doi.org/10.1002/per.2410070103

Cea-D'Ancona, M. A. (2007*). La deriva del cambio familiar. Hacia formas de convivencia más abiertas y democráticas [The drift of family change. Towards more open and democratic forms of coexistence].* Centro de Investigaciones Sociológicas [CIS].

Chittayath, P. (2003). Sobre la familia en la India [On the family in India]. In D. Borobio (Coord.), *Familia e interculturalidad [Family and interculturality]* (pp. 395-400). Publicaciones Universidad Pontifica de Salamanca.

Consejo de Europa. (2006). *Informe explicativo. Recomendación Rec (2006) 19 del Comité de Ministros a los Estados miembros sobre políticas de apoyo a la parentalidad positiva [Recommendation Rec (2006) 19 of the Committee of Ministers to member States on policy to support positive parenting].* https://bit.ly/3YzZdba

Consejo Económico y Social España [CES España]. (2019). *La inmigración en España: efectos y oportunidades [Immigration in Spain: effects and opportunities].* https://www.ces.es/documents/10180/5209150/Inf0219.pdf

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* Plenum Press. https://bit.ly/3YUYlxH

Del Barrio, V., Moreno, C., & López, R. (2001). Evaluación de la agresión y la inestabilidad emocional en niños españoles: su relación con la depresión [Assessment of aggression and emotional instability in Spanish children: its relationship with depression]. *Clínica y Salud, 12* (1), 33-50. https://journals.copmadrid.org/clysa/art/be83ab3ecd0db773eb2dc1b0a17836a1

Donati, P. (2003). *Manual de sociología de la familia [Handbook of the sociology of the family].* Ediciones Universidad de Navarra [EUNSA].

Durkheim, E. (1892). *La famille conjugale [The conjugal family].* Les Éditions de Minuit.

EAPN España. (2021). *El estado de la pobreza. Seguimiento del indicador de pobreza y exclusión social en España 2008-2020 [The state of poverty. Monitoring of the poverty and social exclusion indicator in Spain 2008-2020].* https://bit.ly/3AvmLTk

Eltink, C. F., Chicanelli, A. C., & Lankaster, T. (2024). Afecto familiar y desempeño escolar de los niños de la escuela primaria. Una revisión integradora [Family affection and school performance of children in elementary school. An integrative review]. *Prometeica: Revista de Filosofia y Ciencias*, *29*, 348-364. https://doi.org/10.34024/prometeica.2024.29.16276

Farkas, M. S., & Grolnick, W. S. (2010). Examining the components and concomitants of parental structure in the academic domain. *Motivation and Emotion, 34* (3), 266-279. https://doi.org/10.1007/s11031-010-9176-7

Fernández-Hawrylak, M., Orozco, M. L., & Heras, D. (2016). Familia y migración: las familias transnacionales [Family and migration: The transnational families]. *Familia: Revista de Ciencias y Orientación Familiar*, (53), 87-106. http://bit.ly/3jTImBd

Fundación Foessa, & Cáritas. (2022). *Evolución de la cohesión social y consecuencias de la covid-19 en España [Evolution of social cohesion and the consequences of covid-19 in Spain]*. Fundación Foessa; Cáritas Española Editores. https://bit.ly/3BQyI6A

Grolnick, W., & Pomerantz, E. (2009). Issues and challenges in studying parental control: Toward a new conceptualization. *Child Development Perspectives*, *3* (3), 165-170. https://doi.org/10.1111/j.1750-8606.2009.00099.x

Guo, J. (2003). Sobre la familia en China [On the family in China]. In D. Borobio (Coord.), *Familia e interculturalidad [Family and interculturality]* (pp. 387-390). Publicaciones Universidad Pontifica de Salamanca.

Habiyakare, D. (2003). Familia y cultura africana [African family and culture]. In D. Borobio (Coord.), *Familia e interculturalidad [Family and interculturality]* (pp. 371-380). Publicaciones Universidad Pontifica de Salamanca.

Henao-Agudelo, C., Lalueza, J. L., & Tenorio, M. C. (2016). Valores y prácticas educativas de familias latinoamericanas inmigradas en Barcelona: ¿qué cambia y qué permanece? [Values and educational practices of immigrant families from Latin America to Barcelona: what changes and what stays the same?]. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, *14* (1), 603-615. https://doi.org/10.11600/1692715x.14141090315

Iglesias, J., Rua, A., & Ares, A. (2020). *Un arraigo sobre el alambre. La integración social de la población de origen inmigrante (POI) en España [A rootedness on the wire. The social integration of the population of immigrant origin in Spain]*. Fundación Foessa; Cáritas Española Editores. https://bit.ly/3K3UqdH

Inda-Caro, M., Fernández-García, C. M., & Viñuela-Hernández, M. P. (2023). Parental educational styles from parents' and children's perspectives [Parental educational styles from parents' and children's perspectives]. *Bordón, Revista de Pedagogía*, *75* (3), 85-101. https://doi.org/10.13042/Bordon.2023.94551

Inda-Caro, M., Rodríguez-Menéndez, C., & García-Pérez, O. (2022). Psychometric properties of the Perceived Parental Autonomy Support Scale (P-PASS) in Spanish late-adolescents. *Anuario de Psicología*, *52* (2), 190-200. https://revistes.ub.edu/index.php/Anuario-psicologia/article/view/33615

INE. (2023c). *Estadística del padrón continuo. Población (españoles/extranjeros) por edad (año a año) y sexo [Continuous census statistics. Population (Spaniards/foreigners) by age (year by year) and sex]*. http://bit.ly/3xpPBni

INE. (2023b). *Estadística del padrón continuo. Población (españoles/extranjeros) por edad (grupos quinquenales) y sexo [Continuous census statistics. Population (Spaniards/foreigners) by age (five-year groups) and sex]*. https://bit.ly/3IpoL59

Instituto de Política Familiar. (2021). *Evolución de la familia en España 2021 [Evolution of the family in Spain 2021]*. IPF. https://bit.ly/4aU01wL

Instituto Nacional de Estadística [INE]. (2023a). *Estadística del padrón continuo. Población extranjera por nacionalidad, sexo y año [Continuous census statistics. Foreign population by nationality, sex and year]*. https://bit.ly/3EavPQu

Jorquera, G.; Del Moral, C. (Coord.). (2020). *Familias en riesgo. Análisis de la situación de pobreza en los hogares con hijos e hijas en España [Families at risk. Analysis of the situation of poverty in households with children in Spain]*. Save The Children España. https://bit.ly/3xdx8xV

Llano, J. C. (Dir.) (2023). *El estado de la pobreza. Seguimiento de los indicadores de la Agenda UE2030 (2015-2022) [The state of poverty. Monitoring the EU2030 Agenda indicators (2015-2022)]*. European Anti-Poverty Network (EAPN). https://bit.ly/3VCNE48

Llopis, D., & Llopis, R. (2003). Estilos educativos parentales y relaciones sociales en adolescentes [Parental educational styles and social relationships in adolescents]. *Familia*, 27, 53-70. http://bit.ly/3I1fOh9

Mageau, G. A., Ranger, F., Joussemet, M., Koestner, R., Moreau, E., & Forest, J. (2015). Validation of the Perceived Parental Autonomy Support Scale (P-PASS). *Canadian Journal of Behavioural Science*, 47 (3), 251-262. https://doi.org/10.1037/a0039325

Mejia-Flores, M., Sánchez-Manobanda, K., & Morales-Gómez, M. (2024). La educación afectiva para el desarrollo de la inteligencia emocional [Affective education for the development of emotional intelligence]. *593 Digital Publisher CEIT*, *9* (1-1), 136-147. https://doi.org/10.33386/593dp.2024.1-1.2267

Mestre, V., Pérez-Delgado, E., García Samper, L., & Martí, M. (1997). Educación familiar y desarrollo de la afectividad en los hijos [Family education and the development of affectivity in children]. *Familia: Revista de Ciencias y Orientación Familiar*, (16), 47-64. http://bit.ly/3I1fTRZ

OECD [Organisation for Economic Co-operation and Development]. (2017). Chapter 3. A portrait of family migration in OECD countries. In *International Migration Outlook 2017* (pp. 107-166) https://doi.org/10.1787/migr_outlook-2017-6-en

OECD [Organisation for Economic Co-operation and Development]. (2022). *Evolving family models in Spain: A new national framework for improved support and protection for families.* OECD Publishing. https://doi.org/10.1787/c27e63ab-en

Páez-de la Torre, S. (2020). *La teoría de los capitales de Pierre Bourdieu como modelo de análisis de los procesos de empoderamiento: el caso de los jóvenes originarios del Cono Sur latinoamericano que viven en Catalunya [Pierre Bourdieu's theory of capitals as a model for analysing empowerment processes: the case of young people from the Southern Cone of Latin America living in Catalonia]* [Doctoral dissertation]. Repositorio UdG, Universidad de Girona. http://hdl.handle.net/10256/16347

Pérez-Díaz, V., Rodríguez, J. C., & Sánchez, L. (2001). *La familia española ante la educación de los hijos [The Spanish family and the education of children].* Fundación La Caixa.

Polanco, N. (2003). La familia vista desde Suramérica [The family as seen from South America]. In D. Borobio (Coord.), *Familia e Interculturalidad [Family and interculturality]* (pp. 381-386). Publicaciones Universidad Pontifica de Salamanca.

R Core Team (2022). *The R Project for Statistical Computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Reher, D. S., Cortés, L., González, F., Requena, M., Sánchez, M. I., Sanz, A., & Stanek, M. (2007). *Informe Encuesta Nacional de Inmigrantes (ENI - 2007).* INE; Ministerio de Trabajo en Inmigración; Universidad Complutense; GESP. UNED; https://bit.ly/3bHT2eL

Rocher, G. (1973). *Introducción a la sociología general [Introduction to general sociology].* Herder.

Rodríguez-Meirinhos, A., Antolin-Suárez, L., Brenning, K., Vansteenkiste, M., & Oliva, A. (2020). A bright and a dark path to adolescents' functioning: The role of need satisfaction and need frustration across gender, age, and socioeconomic status. *Journal of Happiness Studies*, *21* (1), 95-116. https://doi.org/10.1007/s10902-018-00072-9

Ryan, R., & Deci, E. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61* (3). https://doi.org/10.1016/j.cedpsych.2020.101860

Schaffer, H. R. (1989). *Interacción y socialización [Interaction and socialisation].* Visor.

Skinner, E., Jonhson, S., & Snyder, T. (2005). Six dimensions of parenting: A motivational model. *Parenting: Science and Practice*, *5* (2), 175-235. https://doi.org/10.1207/s15327922par0502_3

Soenens, B., & Vansteenkiste, M. (2010). A theoretical upgrade of the concept of parental psychological control: Proposing new insights on the basis of self-determination theory. *Developmental Review*, *30* (1), 74-99. https://doi.org/10.1016/j.dr.2009.11.001

Soenens, B., Vansteenkiste, M., & Van Petegem, S. (2015). Let us not throw out the baby with the bathwater: Applying the principle of universalism without uniformity to autonomy-supportive and controlling parenting. *Child Development Perspectives*, *9* (1), 44-49. https://doi.org/10.1111/cdep.12103

Torío-López, S., Peña, J. V., & Inda-Caro, M. (2008). Estilos de educación familiar [Parenting styles]. *Psicothema*, *20* (1), 62-70. https://bit.ly/47NjhL1

Torío-López, S., Rodríguez Menéndez, C., & García-Pérez, O. (2022). La parentalidad positiva y la teoría de la autodeterminación: un avance en la formación parental [Positive parenting and the theory of self-determination: a breakthrough in parenting education]. In S. Rivas y C. Beltramo (Coords.). *Parentalidad positiva. Una mirada a una nueva época [Positive parenting. A look at a new era]* (pp. 47-65). Pirámide.

Tur, A. M. (2003). *Conducta agresiva y prosocial en relación con temperamento y hábitos de crianza en niños y adolescentes [Aggressive and prosocial behaviour in relation to temperament and parenting habits in children and adolescents]* [Doctoral dissertation]. RODERIC, Universitat de València. http://hdl.handle.net/10550/38891

Tur-Porcar, A. M., Doménech, A., & Mestre, V. (2018). Vínculos familiares e inclusión social. Variables predictoras de la conducta prosocial en la infancia [Family linkages and social inclusion. Predictors of prosocial behavior in childhood]. *Anales de Psicología*, *34* (2), 340-348. http://dx.doi.org/10.6018/analesps.34.2.308151

Van der Kaap-Deeder, J., Vansteenkiste, M., Soenens, B., & Mabbe, E. (2017). Children's daily well-being: the role of mothers', teachers', and siblings' autonomy support and psychological control. *Developmental Psychology*, *53* (2), 237-251. https://doi.org/10.1037/dev0000218

Yu, J., Cheah, C. S. L., Hart, C. H., Sun, S., & Olsen, J. A. (2014). Confirming the multidimensionality of psychologically controlling parenting among Chinese-American mothers: Love withdrawal, guilt induction, and shaming. *International Journal of Behavioral Development*, *39* (3), 285-292. https://doi.org/10.1177/0165025414562238

Authors' biographies

María-Elena Rivoir-González. Bachelor's degree in Pedagogy at the Universidad de Oviedo (2018) with the highest rating. Extraordinary Degree Award (2017/2018). Master's degree in Social and Educational Intervention & Research at the Universidad de Oviedo (2021). Education and Psychology PhD student at the Universidad de Oviedo. Predoctoral Researcher at the Department of Educational Sciences at the Universidad de Oviedo (FPU, Ministry of Universities). She teaches at Universidad de Oviedo since 2019 in the Area of Theory and History of Education. She is member of the ASOCED research group (officially recognized by ANECA and the University of Oviedo). Her major research interests involve migration, family and education.

iD https://orcid.org/0000-0002-3434-0826

Carmen-María Fernández-García. Bachelor's degree in Pedagogy at the University of Oviedo (2000) with the highest rating. Extraordinary Degree Award (1999/2000). PhD at the University of Oviedo with the highest rating. Extraordinary PhD Award (2003/2004). Associate Professor at the Department of Educational Sciences at the University of Oviedo, with lectures in the degree of Pedagogy and in the master's degree in Lower Secondary Education, Upper Secondary Education and Vocational Education and Training. She is member of the Spanish Society of Comparative Education, the Spanish Society of Pedagogy and the ASOCED research group (officially recognized by ANECA and the University of

Oviedo). Her major research interests involve gender and education, family education, teaching and teacher education and comparative education.

 https://orcid.org/0000-0001-6314-355X

Susana Torío-López. Bachelor's degree in Philosophy and Arts (speciality: Philosophy and Educational Sciences) at the Pontifical University of Salamanca (1988). PhD at the University of Oviedo (2001), with the highest rating. Extraordinary PhD Award (2002/2003). Associate Professor at the Department of Educational Sciences at the University of Oviedo since 2003, in the Area of Theory and History of Education. Her teaching activity focuses on the sub-area of Theory and History of Education at the Faculty of Teacher Training and Education. She also collaborates as a lecturer in several masters at the University of Oviedo, the University of Granada and the University of Salamanca. She is member and main researcher of the ASOCED Research Group (officially recognized by ANECA and the University of Oviedo). Her major research interests involve family and education or socioeducational intervention with normalized and vulnerable families. Three six-year- research periods assessed by the CNEAI (2014-2019).

 https://orcid.org/0000-0001-5004-2338

Teaching action and altruistic behaviour in Physical Education classes. Predictive analysis applying the 3×2 motivational climate model

Acción docente y comportamiento altruista en clases de Educación Física. Un análisis predictivo desde el modelo de clima motivacional 3×2

María del Carmen FLORES-PIÑERO. Doctoral Student in Educational Sciences. Universidad de Granada, Spain (*maricarmen27@correo.ugr.es*).

Juan GONZÁLEZ-HERNÁNDEZ, PhD. Full Professor. Universidad de Granada, Spain (*jgonzalez@ugr.es*).

Pedro VALDIVIA-MORAL, PhD. Full Professor. Universidad de Granada, Spain (*pvaldivia@ugr.es*).

Abstract:

This study aims to identify the way in which the motivational climate generated by Physical Education (PE) teachers influences engagement in prosocial-altruistic behaviour in a sample of 714 12-to-18-year-old Spaniards adolescents. To this end, the following objectives were proposed: (1) to identify the types of motivational climates established by teachers in PE classes and their impact on facilitating or limiting student engagement in altruistic prosocial behaviours, and (2) to examine the existence of gender differences and describe whether they are also influenced by the climates generated by teachers. A descriptive, non-randomized and relational research design was employed. The students were applied instruments for the 3 × 2 motivational climate and prosocial behavior in relation to their experiences in physical education classes. Findings reveal that an other-oriented climate, whether approach or avoidance oriented, reduces engagement in altruistic prosocial activities in the physical education classroom. Data analysis lead to the conclusion that a more comparative and negatively competitive perception (climates oriented towards how I perform in front of others) decreases engage-

ment in altruistic prosocial behaviour, mainly in boys, promoting instead rivalry and competitive behaviour.

Keywords: prosocial behaviour, motivational climate, prosocial-altruistic behaviour, physical education, teenagers, teachers.

Resumen:

En este estudio, se pretende conocer cómo influye el clima motivacional generado por el profesor de Educación Física (EF) en el desarrollo de la conducta prosocial-altruista. Para ello, los objetivos que se proponen son: (1) conocer qué climas motivacionales construidos por el profesorado en clases de EF facilitan o limitan la aparición de comportamientos prosociales altruistas en su alumnado y (2) describir si existen diferencias de género en el alumnado y si estas se ver influidas de la misma manera por los climas que genera el profesorado. El diseño utilizado ha sido descriptivo, no aleatorizado y relacional, en una muestra de 714 adolescentes españoles de ambos sexos, con edades comprendidas entre los 12 y 18 años. Se han aplicado a los alumnos instrumentos para el clima motivacional 3 × 2 y la conducta prosocial en relación con sus experiencias en las clases de educación física. Los resultados señalan que un clima orientado al otro, ya sea por aproximación, ya sea por evitación, disminuye la posibilidad de que aparezcan conductas prosociales altruistas en las clases de educación física. Tras el análisis de los datos, se concluye que una percepción más comparativa y competitiva (climas orientados a cómo me desenvuelvo frente a los demás) disminuye la conducta prosocial altruista, principalmente en chicos, y fomenta conductas de rivalidad y competitividad hacia el otro.

Palabras clave: comportamiento prosocial, clima motivacional, conducta prosocial-altruista, educación física, adolescentes, profesorado.

1. Introduction

Prosocial behaviour was first studied in the 20th century, specifically from the 1970's, and refers to support behaviours that promote social connection and moral development throughout the life course. Such behaviours promote wellbeing, understanding of social norms and cooperative acts and, in turn, favour self-esteem and positive self-concept (Busching & Krahé, 2020; Crone & Achterberg, 2022; Kolgberg, 2014; Li et al., 2021; Preston & Rew, 2022; Son & Padilla-Walker, 2020).

Social Learning Theory (Bandura, 1982) argues that the internalised and motivational control of behaviour lays the foundations for prosocial tendencies (e.g., altruism) instead of comparative tendencies (e.g., being better than others) (Eisenberg & Spinrad, 2014). The former are linked with the formation of strong and effective personalities (Thielmann et al., 2020), whilst the latter are linked with more unstable traits, depending on interpretations of success/failure and their impact on self-concept: e.g., physical or motor

self-concept regarding Physical Education (PE onwards) or social self-concept in sentimental relationships (González-Hernández & Martínez-Martínez, 2020). On the other hand, prosocial tendencies have also been shown to be protective against engagement in disruptive social behaviours such as violence (Ahmed et al., 2020), aggression (Arbel et al., 2022), narcissism (Kauten & Barry, 2016), and alcohol and substance abuse (Hernández-Serrano et al., 2016), whilst also reducing the development of psychopathologies (Memmott-Elison & Toseeb, 2023) throughout childhood and adolescence.

It is important to highlight that prosocial behaviours include altruistic behaviours but that not all prosocial behaviours are altruistic (Lemos & Richaud, 2013). Padilla-Walker and Carlo (2015) classified prosocial behaviour according to different types, with altruistic tendencies standing out as "prosocial acts performed voluntarily, with a sense of justice and without expecting anything in return either in the short or long term" (p. 8). For a prosocial behaviour to be considered to be altruistic and healthy, it must meet the following parameters: a) help somebody else with a *clean* motivation and not from a selfish standpoint; b) support another person's wellbeing out of empathy; c) be voluntary; d) not expect to receive any immediate extrinsic reward, although an intrinsic reward (driving this behaviour) can be derived, and e) giving more to others than receiving.

Prosocial behaviour can be promoted at any opportunity through school experiences and, specifically, in PE classes (Martí-Vilar

et al., 2019). Acknowledging the educational context as the climate that is promoted in the classroom by teaching staff and perceived by students to facilitate the achievement of both task performance and skill development (Moral-García et al., 2019; Rodrigues et al., 2022; Vasconcellos et al., 2020). Thus, goal achievement assumes a pattern of beliefs pertaining to success, variety of effort and capacity for learning, alongside prosocial feedback between adults and peers (Busching & Krahé, 2020; Cuevas-Campos et al., 2013; Elliot et al., 2013). It mustn't be forgotten that adolescence represents a stage at which numerous changes occur in development, not only at a physical level but, also, at an emotional and psychosocial level. Essential aspects should be promoted, therefore, which connect interpersonal processes (e.g. empathy and altruism) in the educational setting. These will be reflected in the development of prosocial behaviours, in this way improving wellbeing, social adaptation and co-existence (Bisquerra-Alzina & López-Cassà, 2021)

In recent decades, theoretical evolution of the model adapting motivated behaviour to include educational action (Ames, 1992) has focused on the importance of constructing contexts in which class goal structures influence perceptions of competence and ability beyond individual achievement goals (Bardach et al., 2020). Elliot et al. (2011) argue that, in order to feel competent, goal orientations must have an absolute standard ("task focused"), an interpersonal standard ("self-focused") and a more normative and interpersonal standard ("focused on the acts of others"), all of which are assessed according to two

valence indicators (approach and avoidance). These goals are illustrated by the 3×2 model of achievement motivation (Figure 1), which outlines student perceptions of their PE teachers according to perceptions of competence (approach taken to or achievement of desired goals) or incompetence (avoidance or achievement of undesired goals) (Midgley et al., 2000; Murayama et al., 2012).

FIGURE 1. 3×2 theoretical model of achievement goals.
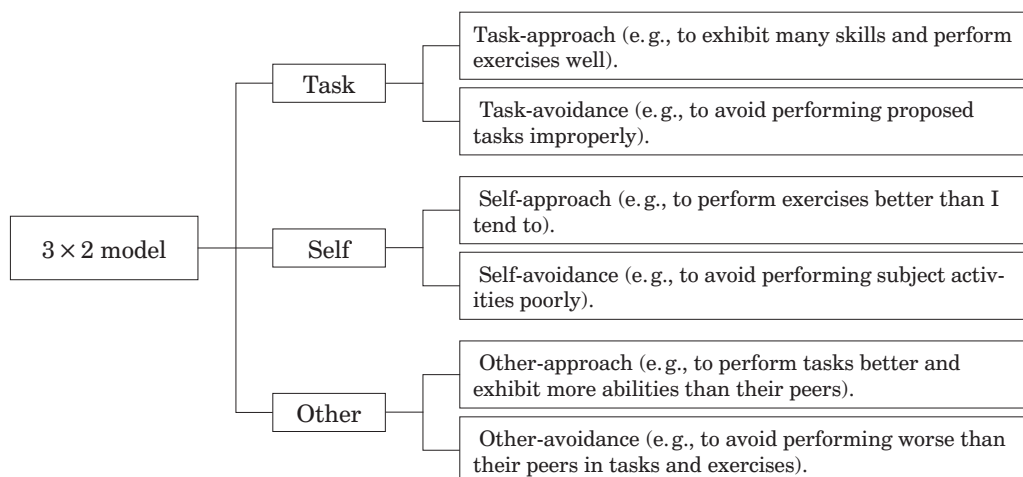
Source: adapted from Elliot et al., 2011.

Finally, the quality of teaching and of the educational system falls on the shoulders of teaching staff. However, sometimes, teachers are not aware of the orientation that they should follow in their training or of the traits that they should possess. Turning attention to PE teaching staff, according to Villaverde-Caramés et al. (2021), the following orientations most stand out: a) to possess pedagogical capacity, subject knowledge, and positive attitude and interest in lifelong training; b) to inspire, feel and transmit values, and to transmit these in class through a pleasant climate; and c) to promote an active and psychosocially healthy lifestyle both inside and outside school.

Following the literature review, a research question is posed around whether the motivational climate generated by PE teachers impacts student engagement in prosocial-altruistic behaviours. In order to address this question, the main aim of the present work is to identify the impact of the motivational climate generated in mandatory secondary education PE classes on the engagement of students in basic altruistic behaviour as a means of healthy development and psychosocial adaptation. To this end, the following specific objectives were proposed: (1) identify whether the motivational climate constructed by teaching staff in PE classes facilitates or impedes student engagement in altruistic prosocial behaviours and (2) examine the existence of gender differences with regards to the first objective and describe whether the influence of the motivational

climate generated by teachers differs as a function of gender.

To this end, a hypothetical model (Figure 2) was designed in accordance with the framework laid out by the 3× 2 model of goal orientation (Elliot et al., 2011). This model argues that task- and self-oriented motivational climates (both approach and avoidance) promote engagement in altruistic prosocial behaviour in PE classes, whilst other-oriented climates

(whether approach or avoidance) impede and reduce engagement in such behaviours. Thus, the following hypotheses were proposed:

- H1: task- and self-oriented motivational climates will positively predict altruistic prosocial behaviour.

- H2: other oriented motivational climates will negatively predict altruistic prosocial behaviour.

FIGURE 2. Hypothetical model of the motivational climate constructed by PE teachers and student altruistic behaviour.

Note: ⟶ Positively predicts altruistic prosocial behaviour.
  – – ⟶ Negatively predicts altruistic prosocial behaviour.
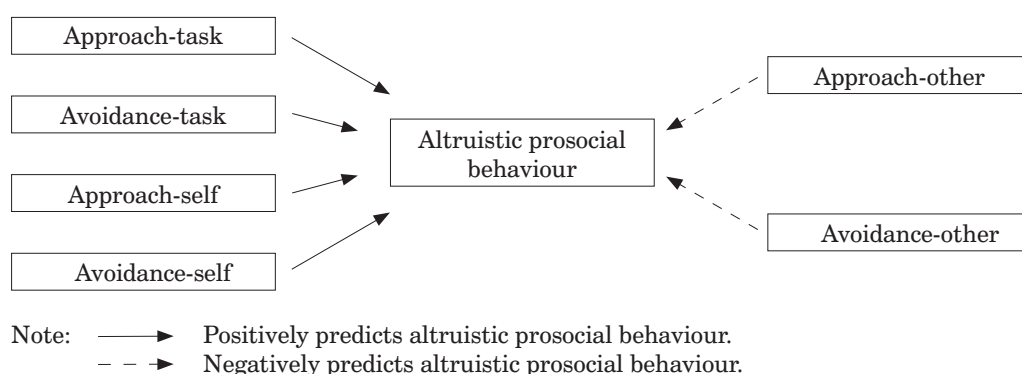
2. Methodology

2.1. Design and procedure

A descriptive, non-randomised and relational design was employed (Gómez, 2019). The study was cross-sectional in nature. In this sense, data were collected at a single time-point between February and May. A final sample of 714 participants was recruited. Participating schools gave permission for researchers to be present in person to administer questionnaires. Questionnaires took between 15 and 20 minutes to complete and were filled out during class form time.

The present research adhered to the principles laid out by the Declaration of Helsinki (1964) and was approved by the ethics committee at the host university (ref: 2296/CEIH/2021). The research protocol pertained to the following steps. First, the researches got in touch with school directors and PE teachers. Secondly, the research aims and research plan were explained and collaboration was requested (receive agreement to participate and informed consent from tutors). Next, access to classrooms was requested for questionnaire administration whilst,

at the same time, instructions were provided for their completion (via electronic devices administered in quite spaces, approximately 15-20 minutes in duration). In order to ensure methodological rigour, data collection was preceded by a preliminary question in which consent was requested and the importance of honesty and sincerity when responding was highlighted. Likewise, study participants were informed that they were free to leave the study at any time and that all data would be kept anonymous and confidential throughout the research process in order to guarantee the protection of personal data.

2.2. Participants

The study population pertained to students enrolled in mandatory secondary education public and subsidised (mixed funding) schools in Andalusia and the Canary Islands. In order to identify the population of potential study participants, electronic contact was made with the various boards of education, sport and culture under the auspices of the Junta de Andalucía. The next step was to make contact with individual secondary schools. Schools that agreed to participate in the present study were sent all necessary information, together with the link to access the survey. A final sample of 714 participants was achieved following removal of incorrectly completed questionnaires. In the Canary Islands, participating schools were selected out of convenience, through a contact who put the research team in touch with schools so that they could be sent study documentation and provided with full details regarding the research process.

The study population comprised 714 adolescents aged between 12 and 18 years (M_{age} = 14.13%; SD_{age} =1.32): 388 males (M_{age} = 14.11% ; SD_{age} = 1.30) and 326 females (M_{age} = 14.15% ; SD_{age} = 1.33) enrolled at public (81.2 %) or subsidised (mixed funding) (18.8 %) mandatory secondary education institutions located in different regions of Andalusia and the Canary Islands (Spain).

2.3. Instruments

Amongst other general variables of interest, the present study gathered data on the sociodemographic characteristics of age (12-18 years), academic year (1st, 2nd, 3rd and 4th year of compulsory secondary education), type of educational institution (public and subsidised) and sex (males and females).

In order to assess motivational climate in PE classes, the 3 x 2 Motivational Climate Questionnaire conceived by Elliot et al. (2011) was employed. This tool has been validated in the Spanish context by Méndez-Giménez et al. (2018). The questionnaire is composed of 18 items that represent the motivational climate structures generated by PE teachers (ego-oriented, task-oriented, other-oriented) and assess student's perceptions of these climates according to direction or valence (approach, avoidance). Internal consistency of this tool when administered within the examined sample produced values of between .66 (task-avoidance) and .78 (other-approach), whilst the Cronbach alpha produced for the overall scale was .87.

In order to measure prosocial behaviour, the Spanish version of the Prosocial Tendencies Measure Revised (Carlo & Randall, 2002; PTM-R) was employed. This produced a reliability index of 0.79. It is composed of 21 items which are rated along a five-point Likert scale ranging from one ("Does not describe me at all") to five ("Describes me very well"). This tool measures six types of prosocial tendencies or motivations, however, for the present study, only items pertaining to altruistic prosocial tendencies were considered. An adjusted reliability index of .71 was produced for the overall scale when applied within the present sample.

2.4. Data analysis

In the present study, version 25 of the statistical software package SPSS (IBM) was used to perform all data analysis. First, preliminary analysis was conducted according to measures of central tendency, missing cases were removed, data distribution assessed (Kolmogorov-Smirnov >.00) and internal consistency of employed measures examined (Cronbach alpha). Next, analysis of categorical variables was performed in order to identify trends in student responses, followed by differential analysis (t test of mean differences), correlation analysis (Pearson) and simple regression analysis (stepwise method) in order to examine the existing relationship between the dependent variable (prosocial behaviour) and independent variables (motivational climate). For each one of these tests, bootstrapping was also performed (sample <5000) as a means of improving the reliability of analyses and achieving acceptable confidence levels (CI >95%).

3. Results

3.1. Differential and correlational analysis

T-test analysis of mean differences as a function of gender revealed that females reported significantly higher scores for altruistic prosocial behaviour [$t_{(712,2)}$ = - 4.682, <.00], avoidance-other motivational climate [$t_{(712,2)}$ = 3.92, <.00] and approach-other motivational climate [$t_{(712,2)}$ = 3.14; <.00].

With regards to linear correlational analysis (Table 1), and controlling for the aforementioned gender differences, outcomes reflect the existence of significant negative relationships between altruistic prosocial behaviour and motivational climates oriented towards avoidance-other (e.g., avoid exhibiting worse performance than others) and towards approach-other (e.g., wanting to exhibit better performance than others). In other words, student perceptions of competitiveness and rivalry within their motivational climate meant that they tended to engage less in altruistic behaviour.

In addition, it was also observed, although correlation indices were small, that increasing age was associated with decreasing perceptions of an approach-ego orientation (-.08*) and increasing perceptions of an approach-other orientation (.08*). This reveals greater tendencies in these students towards rivalry.

TABLE 1. Correlational analysis adjusted for gender.

	1	2	3	4	5	6	7	8
Age	1	-.033	-.087[*]	.080[*]	.026	-.054	-.052	.070
Approach-task		1	.674[**]	.154[**]	.246[**]	.404[**]	.517[**]	-.030
Approach-ego			1	.146[**]	.222[**]	.458[**]	.545[**]	-.041
Approach-other				1	.757[**]	.278[**]	.136[**]	-.182[**]
Avoidance-other					1	.402[**]	.334[**]	-.177[**]
Avoidance-ego						1	.618[**]	.026
Avoidance-task							1	.019
Prosocial altruism								1

Note: [**]significance at the level .01 (two-tailed); [*]significance at the level .05 (two-tailed).
1 = age; 2 = approach-task; 3 = approach-ego; 4 = approach-other; 5 = avoid-other; 6 = avoid-ego; 7 = avoid-task; 8 = pro-altruist.

3.2. Multiple regression analysis

In order to enable more in-depth analysis of the relationship between variables in a PE context, multiple regression (following a stepwise approach) was performed, which resulted in a final model that explained 27.3% of the variability found in student engagement in prosocial behaviour (DV) according to the motivational climates created by PE teachers (IVs) ($F_{(707,7)}$ = 26.04; R^2 = .273; p <.05) (Table 2).

Controlling for gender, predictive analysis revealed that students who do not perceive an other-oriented motivational climate (which entails students comparing themselves with their peers, promoting competitive behaviours and rivalry) tend to engage in more altruistic prosocial behaviour. In addition, suitable perceptions of the duties performed by teachers with regards to avoidance-ego also reduce comparisons with others. As a result, students are able to become less self-critical (mainly destructive), exhibit fewer tendencies with regards to inferiority/superiority over others and further support engagement in altruistic behaviour. In the same way, age also emerged as a variable that has an impact on student engagement in altruistic prosocial behaviour in PE classes. As students grow and mature, greater emphasis is placed on forming positive connections with others, reducing competitive outlooks and leaving rivalries with peers to one side.

TABLE 2. Predictive analysis of altruistic prosocial tendencies adjusted for gender.

Model	R	R^2	F	Independent variable	
DV: altruistic prosocial behaviour				**Variable**	**β**
1	.182(a)	.033	24.41**		
				Approach-other	-.182***
2	.201(b)	.040	14.97**		
				Approach-other	-.189***
				Age	.086*
3	.219 (c)	.048	11.91**		
				Approach-other	-.215***
				Age	.093*
				Avoidance-ego	.090*
4	.235 (d)	.055	10.36**		
				Approach-other	-.118*
				Age	.090*
				Avoidance-ego	.119*
				Avoidance-other	-.138*

Note: **significance at the level .01 (two-tailed); *significance at the level .05 (two-tailed). Excluded variables: approach-task and avoidance-task

4. Discussion

The aim of the present study was to identify the importance of the motivational climate generated by teaching staff in PE classes with regards to student engagement in altruistic prosocial behaviour. In order to develop and plan teaching strategies in PE (and in any other academic subject), in may be useful for teachers to consider the combined fit of strategies in order to deliver experiences based on the three types of climates proposed by the 3 × 2 model (task, ego and other). From a pedagogical standpoint, these have important repercussions on student perceptions.

Present findings confirm H1: task- and other-oriented motivational climates would positively predict altruistic prosocial behaviour. Indeed, significant positive relationships emerged between altruistic prosocial behaviours and motivational climates oriented towards self (e.g., striving to improve) and task (e.g., effort towards learning).

In accordance with that outlined by García-Romero (2015), strong correlations were produced between goal orientations proposed by the 3 × 2 model of motivational climate (self-approach and self-avoidance; task-approach and task-avoidance;

other-approach and other-avoidance), with the strongest relationships pertaining to approach-task and approach-self goals. In the same way, the achievement motivation model conceived by Nicholls (1989), which shares similar foundations and has been adapted to the school context (Ames, 1992), describes that motivational climates oriented towards cooperation and helping others (prosocial tendencies) prevail when task-oriented motivational climates are perceived and are diminished when ego- or competence-oriented climates are perceived. This is supported by research conducted by Bardach et al. (2020) and Martí-Vilar et al. (2019), in which the 2 x 2 model was used.

In the case of H2, motivational climates oriented towards the other negatively predict altruistic prosocial behaviour. Findings indicate that an other-oriented climate, whether characterised by approach or avoidance, decreases engagement in altruistic prosocial behaviour. Findings coincide with those reported by other similar studies performed in different cultural contexts (Baños, 2021). Moreno et al. (2005) conducted a study with adolescents in which participating students were found to perceive a competence-involved climate oriented towards the other and, at the same time, exhibited disruptive behaviour. Moral et al. (2021) discovered that adolescents tended more towards a task-approach-oriented motivational climate, with this being even more the case amongst boys than girls. Nonetheless, outcomes reported by Moral-García et al. (2019) do not coincide with those reported in this study, given that all

students, regardless of gender, agreed that PE teachers instilled a climate oriented towards skill improvement and task performance in class (20.9% of boys and 17.6% of girls).

Further, tendencies towards altruistic behaviours and outcomes pertaining to motivational climate orientations have been related with age. With increasing age, adolescents tend more towards striving for better task performance in comparison with their peers, with positive correlations with age and an other-approach-oriented motivational climate, and a negative correlation with an ego-approach-oriented motivational climate. Jerez and Cabrera-Fernández (2021) conducted a study with university students and reported a positive correlation with a task climate and a negative correlation with an ego climate. In the same way, present findings and those reported by the aforementioned research concur, given that age was also found to be an influencing factor with regards to student engagement in altruistic prosocial behaviour in PE classes. As adolescents grow in age and maturity, greater emphasis is placed on forming positive connections with others, diminishing competitive outlooks and leaving rivalries with others to one side.

The present study employed the 3 x 2 model which is derived from a prior model (Elliot & McGregor, 2001), in which motivational climates are divided into two types, namely, ego and task. In this case, an ego-oriented climate pertains to competitiveness with oneself, as one strives to improve and perform tasks better that

the previous day. In contrast, an other-oriented climate pertains to competition with others. As clearly demonstrated in the present study, students who perceive an approach-other or avoidance-other motivational climate exhibit less altruistic prosocial behaviour. This concurs with findings reported in the two aforementioned research studies.

With regards to gender differences, it is concluded that girls show higher indicators of altruistic tendencies than boys despite them also exhibited higher indicators of other-oriented climates (approach-avoidance). Existing scientific literature highlights that both types of climates promote the behaviours of rivalry and competition between peers, in which adolescents (mainly boys) strive to be better than everybody else and avoid any type of failure during task performance (Johnson & Arduiz, 2021; Ruiz-Juan et al., 2011; Rodrigues et al., 2022). In addition, the data are inspiring to know that, in a large sample, girls compete in comparison with others (without going into whether with other girls or also with boys), arousing interest to further explore the reasons for such results.

These findings do not coincide with those reported by various other studies conducted in recent years, in which males reported engaging in more altruistic prosocial behaviour (Alandette & Hoyos 2009; Pastor et al., 2024).

Whilst González-Hernández and Martínez-Martínez (2020) concluded that girls engaged in more prosocial behaviour than boys (e.g. greater social ability regarding helping others). Kamas and Preston (2021) argued that prosocial behaviour may differ in boys and girls as a function of their personality and level of empathy.

Findings of the present study with respect to adolescent prosocial tendencies coincide with the meta-analysis performed by Xiao et al. (2019). This review focused on describing gender differences in different dimensions of prosocial behaviour as measured by the Prosocial Tendencies Measure (PTM) (Carlo & Randall, 2002). This same study also uncovered that girls reported higher scores when it comes to emotional altruistic prosocial behaviour, whilst boys reported higher scores with regards to public prosocial behaviour.

Whilst there is a scarcity of research linking achievement contexts and engagement in altruistic behaviour, it can be deduced that motivational climates constructed by teachers in PE classes either support or impinge engagement in altruistic prosocial behaviour within their students. This is in accordance with the outcomes produced in the present study. Thus, further studies that delve deeper into this line of research are of vital importance in order to guide teaching staff in terms of both awareness raising, performance and the improvement of personal teaching tools, and enriching teaching strategies that encourage prosocial behaviours. This will be crucial to counteract tendencies towards comparisons in some students in PE classes. Such research should entail the design of longitudinal proposals than allow observation of the evolution of altruistic behaviours encouraged by teachers

and peers in PE classrooms, and the inclusion of other study variables that provide more information complement previously reported data. A further possibility is to assess whether the influence of general characteristics (e.g., teacher's gender, type of educational institution, differences according to age and academic year) may vary when subjected to the combined influence of the factors observed and examined in the present work.

This being said, the present work is not exempt of limitations with regards to its initial proposal, structural limitations (e.g., the delay in securing permission and consent from participating school directors and tutors lead to a considerable delay in data collection) and theoretical foundation (e.g., the infancy of research examining altruistic prosocial behaviour in PE classes framed by the 3×2 model of achievement motivation). In addition, the cross-sectional nature of the research methodology urges careful and cautious consideration of the produced outcomes, which only have scientific value in the context and time-frame in which they were collected, making it necessary to replicate these findings. Thus, a proposed future perspective is to continue to work and investigate within the field of prosocial behaviour and its related dimensions and motivational climate in the subject of PE with populations of adolescent students.

5. Conclusions

Following data analysis and addressing the aims of the present research, it is concluded that student perceptions of an other-oriented (approach-avoidance) motivational climate decreases altruistic prosocial behaviour, mainly in boys, promoting rivalry and competitiveness towards others.

In this sense, reducing comparative and competitive orientations in educational climates requires teachers to play an important role when it comes to supporting social relationships in their students with the dynamics of PE classes (e.g., by reinforcing the acquisition of positive stereotypes and promoting equality between the sexes). Facilitating prosocial feedback during the teaching-learning process leads teaching practice to shine a brighter light on the tasks, responsibilities and speed of progression inherent to the activities performed to achieve educational goals, and social value of the activity. In this way, reinforcing and improving the social-emotional climate in the classroom turns lessons into social experiences that are of great importance and experientiality, promoting values and supporting the development of a positive environment.

Although only few studies have examined the relationships between these variables, it is possible to conclude that the findings produced in the present study are in accordance with those reported by other research that does exist. In addition, the influence of motivational climate and gender was observed, with boys exhibiting a greater tendency towards competitiveness and rivalry (other-oriented climate) and girls exhibiting a greater tendency towards prosocial behaviour.

Throughout conducting the present research, a number of limitations were

detected that should be considered in future research with the aim of improving and continuing to advance by providing meaningful data that is of interest to this field of study. One of the main limitations is the scarcity of scientific literature examining the combination of the variables of interest, making it impossible to compare present outcomes with those reported in other studies. This being said, at the same time, this opens up the future perspective of continuing to examine this line of research. Further, in opting for a cross-sectional approach, it is difficult to determine the temporal (and causal) sequence of events or determine whether a given variable had a direct influence on another. For this reason, it is important to keep in mind that outcomes from cross-sectional designs may be influenced by factors which pertain to both the interpretation of study participants and, also, the interpretation of findings overall (e. g., recall, measurement and selection bias).

To this end, a number of future perspectives are proposed. Firstly, examine larger samples and extend the present research into other educational institutions. Secondly, compare outcomes between the Canary Islands and Andalusia. Thirdly, design and deliver training programs for PE teachers on the influence of the motivational climate they promote in their classes and student psychosocial development.

6. Authors' contributions

María del Carmen Flores-Piñero: Conceptualisation; Data curation; Methodology; Project administration; Resources; Supervision; Writing (original draft); Writing (review & editing).

Pedro Valdivia-Moral: Methodology; Project administration; Resources; Supervision; Validation; Visualisation; Writing (original draft); Writing (review & editing).

Juan González-Hernández: Data curation; Formal analysis; Methodology; Project administration; Supervision; Validation; Visualisation; Writing (original draft); Writing (review & editing).

References

Ahmed, S., Foulkes, L., Leung, J. T., Griffin, C., Sakhardande, A., Bennett, M., & Blakemore, S. J. (2020). Susceptibility to prosocial and antisocial influence in adolescence. *Journal of Adolescence*, *84* (1), 56-68. https://doi.org/10.1016/j.adolescence.2020.07.012

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84* (3), 261-271. https://doi.org/10.1037/0022-0663.84.3.261

Alandette, Y., & Hoyos, O. (2009). Representaciones mentales sobre los tipos de agresión en escolares [Mental representations of types of aggression in school children]. *Psicología desde el Caribe*, (24), 1-25.

Arbel, R., Maciejewski, D. F., Ben-Yehuda, M., Shnaider, S., Benari, B., & Benita, M. (2022). Prosocial behavior and aggression in the daily school lives of early adolescents. *Journal of Youth and Adolescence*, *51* (8), 1636-1652. https://doi.org/10.1007/s10964-022-01616-2

Bandura, A. (1982). *Teoría del aprendizaje social [Social learning theory]*. Espasa-Calpe.

Baños, R. (2021). Clima motivacional y conductas disruptivas en Educación Física en estudiantes españoles y mexicanos de educación secundaria [Motivational climate and disruptive behaviors in Physical Education in Spanish and Mexican high school students]. *Journal of Sport and Health Research*, *13* (1), 1-12.

Bardach, L., Oczlon, S., Pietschnig, J., & Lüftenegger, M. (2020). Has achievement goal theory been right? A meta-analysis of the relation between goal structures and personal achievement goals. *Journal of Educational Psychology*, *112* (6), 1197-1220. https://doi.org/10.1037/edu0000419

Bisquerra-Alzina, R., & López-Cassà, E. (2021). The intelligent cultivation of moral emotions in adolescence. **Revista Española de Pedagogía**, *79* (278), 103-113. https://doi.org/10.22550/REP79-1-2021-09

Busching, R., & Krahé, B. (2020). With a little help from their peers: The impact of classmates on adolescents' development of prosocial behavior. *Journal of Youth and Adolescence*, *49*, 1849-1863. https://doi.org/10.1007/s10964-020-01260-8

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescence. *Journal of Youth and Adolescence*, *31* (1), 31-44. https://doi.org/10.1023/A:1014033032440

Crone, E. A., & Achterberg, M. (2022). Prosocial development in adolescence. *Current Opinion in Psychology*, *44*, 220-225. https://doi.org/10.1016/j.copsyc.2021.09.020

Cuevas-Campos, R., García-Calvo, T., & Contreras, O. (2013). Perfiles motivacionales en Educación Física: una aproximación desde la teoría de las metas de logro 2 × 2 [Motivational profiles in Physical Education: An approach from the 2 × 2 achievement goals theory]. *Anales de Psicología*, *29* (3), 685-692. https://doi.org/10.6018/analesps.29.3.175821

Eisenberg, N., & Spinrad, T. L. (2014). Multidimensionality of prosocial behavior: Rethinking the conceptualization and development of prosocial behavior. In L. M. Padilla-Walker, & G. Carlo (Eds.), *Prosocial development: A multidimensional approach* (pp. 17-39). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199964772.003.0002

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80* (3), 501-519. https://doi.org/10.1037/0022-3514.80.3.501

Elliot, A. J., Eder, A. B., & Harmon-Jones, E. (2013). Approach–avoidance motivation and emotion: Convergence and divergence. *Emotion Review*, *5* (3), 308-311. https://doi.org/10.1177/1754073913477517

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3 × 2 achievement goal model. *Journal of Educational Psychology*, *103* (3), 632-648. https://doi.org/10.1037/a0023952

García-Romero, C. (2015). Relación entre las metas de logro 3 × 2 y la competencia percibida en los estudiantes de Educación Física [Relationship between the 3 × 2 achievement goals and perceived competence in Physical Education students]. *Sportis: Revista Técnico-Científica del Deporte Escolar, Educación Física y Psicomotricidad*, *1* (3), 293-310.

Gómez, S. (2019). *Metodología de la investigación [Research methodology]*. Red Tercer Milenio.

González-Hernández, J., & Martínez-Martínez, F. D. (2020). Prosociality and socialization difficulties in adolescence. Influences according to sex and sport practice. *Revista de Psicologia del Deporte*, *29* (2), 117-124.

Hernández-Serrano, O., Espada, J. P., & Guillén-Riquelme, A. (2016). Relación entre conducta prosocial, resolución de problemas y consumo de drogas en adolescentes [Relationship between prosocial behaviour, problem-solving and drug use in adolescents]. *Anales de Psicología*, *32* (2), 609-616. https://doi.org/10.6018/analesps.32.2.204941

Jerez, P. & Cabrera-Fernández, A. (2021). Clima motivacional percibido en el deporte y su asociación con los niveles de resiliencia en estudiantes universitarios [Relationship of the prosocial behaviour, the problem-solving skills and the use of drugs amongst adolescents]. *Journal of Sport and Health Research*, *13* (3), 505-514.

Johnson, M., & Arduiz, N. (2021). Emociones del profesorado de Educación Física: revisión narrativa (2010-2020) [Physical education teacher emotions: Narrative review (2010-2020)]. *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, *39*, 910-914.

Kamas, L., & Preston, A. (2021). Empathy, gender, and prosocial behavior. *Journal of Behavioral and Experimental Economics*, *92*, 101654. https://doi.org/10.1016/j.socec.2020.101654

Kauten, R., & Barry, C. (2016). Adolescent narcissism and its association with different indices of prosocial behavior. *Journal of Research in Personality*, *60*, 36-45. https://doi.org/10.1016/j.jrp.2015.11.004

Lemos, V. N., & Richaud, M. C. (2013). Childhood prosocial behavior in the school environment. In A. Castro (Ed.), *Positive psychology in Latin America* (pp. 179-193). Springer

Li, J., Yao, M., & Liu, H. (2021). From social support to adolescents' subjective well-being: The mediating role of emotion regulation and prosocial behavior and gender difference. *Child Indicators Research*, *14*, 77-93. https://doi.org/10.1007/s12187-020-09755-3

Martí-Vilar, M., Corell-García, L., & Merino-Soto, C. (2019). Systematic review of prosocial behavior measures. *Revista de Psicología*, *37* (1), 349-377. https://doi.org/10.18800/psico.201901.012

Memmott-Elison, M. K., & Toseeb, U. (2023). Prosocial behavior and psychopathology: An 11-year longitudinal study of inter-and intraindividual reciprocal relations across childhood and adolescence. *Development and Psychopathology*, *35* (4), 1982-1996. https://doi.org/10.1017/S0954579422000657

Méndez-Giménez, A., Cecchini-Estrada, J. A., & Fernández-Río, J. (2018). Validación del cuestionario de clima motivacional 3 × 2 en Educación Física [Validity of the 3 × 2 motivational climate questionnaire for Physical Education]. *Universitas Psychologica*, *17* (1), 135-148. https://doi.org/10.11144/Javeriana.upsy17-1.vccm

Midgley, C., Maehr, M., Hruda, L., Anderman, E., Anderman, L., Freeman, K. & Gheen, M. (2000). *Manual for the patterns of adaptive learning scales*. University of Michigan.

Moral-García, J. E., Rodríguez, J. N., García-Cantó, E., Pérez-Soto, J. J., Rosa-Guillamón, A., López-García, S., Amatria-Jimenez, M. & Dios, R. (2019). Perception of adolescents on physical education classes according to the physical activity level, weight status, gender and age. *Journal of Physical Education & Sport*, *19*, 40-48. https://doi.org./10.7752/jpes.2019.s1007

Moreno, J. A., Alonso, N., Martínez, C. & Cervelló, E. (2005). Motivación, disciplina, coeducación y estado de *flow* en educación física: diferencias según la satisfacción, la práctica deportiva y la frecuencia de práctica [Motivation, discipline, co-education and flow status in physical education: Differences according to satisfaction, sport practice and frequency of practice]. *Cuadernos de Psicología del Deporte*, *5* (1-2), 225-243.

Murayama, K., Elliot, A. J., & Friedman, R. (2012). Achievement goals. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 191-207). Oxford University Press.

Nicholls, J. G. (1989). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91* (3), 328-346. https://doi.org/10.1037/0033-295X.91.3.328

Padilla-Walker, L. M., & Carlo, G. (Eds.) (2015). *Prosocial development: A multidimensional approach*. Oxford University Press.

Pastor, Y., Pérez-Torres, V., Thomas-Currás, H., Lobato-Rincón, L. L., López-Sáez, M. A. & García, A. (2024). A study of the influence of altruism, social responsibility, reciprocity, and the subjective norm on online prosocial behavior in adolescence. *Computers in Human Behavior*, *154*, 1-10. https://doi.org/10.1016/j.chb.2024.108156

Rodrigues, F., Monteiro, D., Teixeira, D., & Cid, L. (2022). Understanding motivational climates in Physical Education classes: How students perceive learning and performance-oriented climates by teachers and peers. *Current Psychology*, *41* (8), 5298-5306. https://doi.org/10.1007/s12144-020-01047-x

Ruiz-Juan, F., Piéron, M. & Zamarripa, J. (2011). Spanish version of the Task and Ego Orientation in Sport Questionnaire (TEOSQ) adapted to Physical Education. *Studies in Psychology*, *32* (2), 179-193. https://doi.org/10.1174/021093911795978162

Son, D., & Padilla-Walker, L. M. (2020). Happy helpers: A multidimensional and mixed-method approach to prosocial behavior and its effects on friendship quality, mental health, and well-being during adolescence. *Journal of Happiness Studies*, *21* (5), 1705-1723. https://doi.org/10.1007/s10902-019-00154-2

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, *146* (1), 30-90. https://doi.org/10.1037/bul0000217

Vasconcellos, D., Parker, P. D., Hilland, T., Cinelli, R., Owen, K. B., Kapsal, N., Lee, J., Antczak, D., Ntoumanis, N., Ryan, R. M., & Lonsdale, C. (2020). Self-determination theory applied to physical education: A systematic review and meta-analysis. *Journal of Educational Psychology*, *112* (7), 1444-1469. https://doi.org/10.1037/edu0000420

Villaverde-Caramés, E., Fernández-Villarino, M., Toja-Reboredo, B., & González-Valeiro, M. (2021). Revisión de la literatura sobre las características que definen a un buen docente de Educación Física: consideraciones desde la formación del profesorado [A literature review of the characteristics that define a good physical educationteacher: Considerations from teacher training]. *Retos*, *41*, 471-479. https://doi.org/10.47197/retos.v0i41.84421

Xiao, S. X., Hashi, E. C., Korous, K. M., & Eisenberg, N. (2019). Gender differences across multiple types of prosocial behavior in adolescence: A meta-analysis of the prosocial tendency measure-revised (PTM-R). *Journal of Adolescence*, *77*, 41-58. https://doi.org/10.1016/j.adolescence.2019.09.003

Authors' biographies

María del Carmen Flores-Piñero. PhD student in the Faculty of Education Sciences at the University of Granada. Graduated in Pedagogy and with an official postgraduate master's qualification in Psychopedagogical Intervention and a master's degree in Teacher Training in Secondary Education, Languages and Vocational Training (speciality: educational guidance) from the University of Granada. She has been working as an early childhood educator since 2018.

iD https://orcid.org/0000-0001-7469-0233

Juan González-Hernández. Lecturer within the Faculty of Psychology at the University of Granada. He teaches on the postgraduate Teacher Training in Secondary Education, Languages and Vocational Training at the universities of Granada and Murcia. He is a trained sports psychologist with extensive experience in both intervention and research in grassroots sports and the impact of their imbalances on the health of athletes.

iD https://orcid.org/0000-0002-6640-0352

Pedro Valdivia-Moral. Vice dean of placements and lecturer within the Faculty of Educational Sciences at the University of Granada. He delivers teaching on the degree of Teaching in Primary Education and Teaching in Early Childhood Education, in addition to the master's in Teacher Training in Secondary Education, Languages and Training Professionals at the universities of Granada and Jaén. Has worked as a compulsory secondary education teacher at various institutions under the auspices of the Ministry of Education belonging to the Government of Andalusia.

iD https://orcid.org/0000-0002-1905-3247

Binaural music and educational attainment: Evidence from longitudinal and quasi-experimental data

Música binaural y logro educativo: evidencia con datos longitudinales y cuasiexperimentales

Edgar Demetrio TOVAR-GARCÍA, PhD. 1. Full Professor. Universidad Panamericana, Guadalajara, México (*dtovar@up.edu.mx*). 2. Research Fellow. HSE University, Russia (*etovar@hse.ru*).

Abstract:

This article investigates the impact of background binaural music on educational attainment within the context of a natural experiment. In a university business school, 73 undergraduate students participated in this study, providing a maximum of 707 observations at 16 points in time. The students were divided into three groups (Microeconomics A and B, and International Economics) and were further categorized into control and experimental groups at different stages of the courses. All students attended classes from January to June 2021, with the addition of binaural music as a treatment in either the first or second part of the courses (divided into three parts). After each lecture, students completed a multiple-choice quiz designed to assess their learning during the lecture. Utilizing t-tests for mean comparisons and panel data econometric techniques, the study found that the scores of the experimental groups were significantly higher than those of the control groups. Overall, binaural music was observed to increase quiz scores by 8 to 20 basis points.

Keywords: economics, educational attainment, binaural music, natural experiment, longitudinal studies.

Resumen:

Este artículo investiga el impacto de la música binaural de fondo en el logro educativo dentro del contexto de un experimento natural. En el estudio, participaron 73 estudiantes de una escuela de negocios universitaria, con un máximo de 707 observaciones en 16 puntos en el tiempo. Los estudiantes fueron divididos en tres grupos (Microeconomía A y

B, y Economía Internacional) y categorizados, además, como grupos de control y experimentales en diferentes etapas de los cursos. Todos los estudiantes asistieron a clases desde enero hasta junio de 2021, con la adición de música binaural como tratamiento en la primera o segunda parte de los cursos (divididos en tres partes). Después de cada clase, los estudiantes completaban un examen rápido *(quiz)* de opción múltiple diseñado para evaluar su aprendizaje durante la clase. Mediante pruebas *t* para comparaciones de medias y

técnicas econométricas de datos de panel, se descubrió que las puntuaciones de los grupos experimentales fueron significativamente más altas que las de los grupos de control. En general, se observó que la música binaural aumentaba las puntuaciones de los exámenes rápidos entre 8 y 20 puntos básicos.

Palabras clave: economía, logro educativo, música binaural, experimento natural, estudios longitudinales.

1. Introduction

Gaining and keeping students' attention are major concerns when teaching economic theory at the university level, as should be the case for many other subjects. It is well known that millennial and post-millennial generations (digital natives) have difficulties paying attention during lectures (Koponen, 2019; Sharma et al., 2020). This challenge became particularly complex during and shortly after Covid-19, especially with the current hybrid teaching models, which involve both in-person and online students (Keržič et al., 2021; Nieto-Escamez & Roldán-Tapia, 2021). To attract attention, various pedagogical strategies are useful, such as joking, asking interesting questions, team activities, exercises, and quizzes (Briscoe et al., 2022; Castillo-Montoya, 2019; Halawa et al., 2020; Schroeder et al., 2007). In this context, music has been highlighted as a useful tool for concentrating and studying, even in the classroom (Dosseville et al., 2012).

Music can evoke emotions and enhance mood, contributing to better concentration, motivation, and overall well-being, potentially impacting academic achievement (Mega et al., 2014; Pekrun et al., 2002). In addition, listening to music and/or participating in musical activities can reduce stress and anxiety levels, influencing student's ability to focus and perform well academically (Algailani et al., 2023; McBride & Greeson, 2023; Ye et al., 2019). Particularly, instrumental classical music has been shown to reduce stress (Pelletier, 2004) and is useful for solving mathematical tasks (Pavlyugina et al., 2012). Importantly, empirical evidence suggests its positive impact on lecturing (students learn better) (Dosseville et al., 2012).

In educational research, the relationship between music and cognition is a subject of growing interest (Holmes, 2021). Nevertheless, the evidence on the impact of music on cognitive performance is mixed and varies depending on

the type of music, the task being performed, and the population under study (Cheah et al., 2022). Overall, the literature suggests that music has positive effects on different aspects of academic performance (Antony et al., 2018; Ishiguro et al., 2023).

Having said this, recently, binaural music has gained popularity in social networks, arguing its efficiency and efficacy in improving memory and concentration for studying. As such, binaural beats have reached a broader audience by incorporating this technology into music and claiming their usefulness for enhancing focus (Filimon, 2010; Rahman et al., 2021; Sharma et al., 2017). Overall, binaural beats may impact emotional states, physiological responses, and cognitive functions, making binaural music (the masking of binaural beats with music) a relevant subject of study due to its potential effects on memory, attention, stress, anxiety, pain perception, and many other aspects (Garcia-Argibay et al., 2019). Nonetheless, binaural beats have also been criticised due to their potential to cause discomfort (Rahman et al., 2021).

As new technologies and pedagogical tools emerge, the use of auditory stimuli on cognitive performance has also become a subject of growing interest (Cheah et al., 2022; Dosseville et al., 2012; Hallam & Price, 1998; Hickey et al., 2020). The current research contributes to this literature by examining the impact of background binaural music on educational attainment, assessed by multiple-choice quiz scores administered after lectures at different time points. Importantly, this research is based on a natural experiment, differing from the self-reported survey research that characterizes much of the existing literature. For this research, the covid-19 pandemic and the need to innovate in the teaching-learning process allowed the use of binaural music in class, providing longitudinal and quasi-experimental data, whose analysis favours the positive impact of binaural music on educational outcomes. Therefore, this research increases credibility, allowing for an in-depth exploration of the potential influence of binaural music on cognitive performance in real-world educational settings.

The rest of the article is structured as follows. Section 2 presents a literature review on education, music, and binaural beats. Section 3 presents the method: participants, materials, procedure, and data analysis (descriptive statistics, *t*-test for mean comparisons, and regression models with panel data). Section 4 presents the main results and section 5 discusses them, underlining limitations and future research lines. Finally, it concludes with some recommendations.

2. Literature review

Music education is associated with creativity (Burnard & Younker, 2004; Gershon & Ben-Horin, 2014). The improvisational nature of musical activities and the exploration of various

musical elements contribute to the development of creative thinking skills positively impacting problem-solving and critical thinking skills, which is necessary for academic success (Albar & Southcott, 2021; Burnard & Younker, 2004). In addition, music education is linked to socioemotional development (Blasco-Magraner et al., 2022; Ilari, 2020), particularly supporting child development, where collaborative musical experiences, such as group performances, help in developing interpersonal skills, teamwork, and empathy (Ilari et al., 2019). These social skills are relevant to students' overall development and can influence academic success. Likewise, the learning of a musical instrument is linked to cognitive skills, including memory, attention, and spatial-temporal abilities (Rose et al., 2019; Suárez et al., 2016).

In the literature, the positive association between music education and mathematical proficiency is also emphasized (Gillanders & Casal, 2020; Guhn et al., 2019; Jaschke et al., 2013). Music implies understanding patterns, rhythm, and proportions, skills that may be transferred to mathematical reasoning. In general, students who participate in music education exhibit good mathematical skills (Baker et al., 2023; Brazhnikova, 2016).

The emotional and psychological impact of music on students is another important area of study (Algailani et al., 2023; Cheah et al., 2022; McBride & Greeson, 2023; Mega et al., 2014;

Pekrun et al., 2002; Ye et al., 2019). Music affects emotions and exposure to music is associated with improved mood and reduced stress and anxiety levels, positively influencing student's ability to focus, engage, and perform well academically (Mega et al., 2014; Pekrun et al., 2002).

Nevertheless, the relationship between music and academic achievement is complex, with several findings under debate (Cheah et al., 2022). For instance, the Mozart effect (Rauscher et al., 1993) refers to findings indicating that listening to Mozart's music, particularly the *Sonata for two pianos in D major, K. 448*, composed in 1781, can induce short-term improvement in certain types of mental tasks, known as *spatial-temporal reasoning* (Beauvais, 2015; Stough et al., 1994). The Mozart effect is not limited to classical music or the music of the Viennese composer, and it is not without criticism (Nantais & Schellenberg, 1999; Thompson et al., 2001; Waterhouse, 2006).

Overall, integrating music into educational curricula can support student interest and engagement (Algailani et al., 2023; Baker et al., 2023; Hallam & Price, 1998; Ishiguro et al., 2023). The motivation produced by a positive musical experience has positive indirect effects on educational achievement by promoting an active and engaged learning environment. Dosseville et al. (2012) present evidence of the positive impact of classical music on the academic performance of undergraduate students

during learning. This genre of music (its tempo, tonality, modes, etc.) appears to positively influence the learning environment and the affective states of students, thereby helping their comprehension of lectures and improving overall learning outcomes.

In general, the literature supports the positive impact of music on educational outcomes, but potential challenges should be acknowledged, including issues related to resource allocation, student interest, and the integration of music into diverse educational settings. In this context, binaural music has recently emerged and calls for the understanding of its potential benefits and challenges in educational settings, which is crucial for educators, policymakers, and researchers.

2.1. Binaural beats

Binaural beats are two sounds of slightly different frequencies, one for each ear, processed by the brain creating the sense (illusion) of a third sound (Filimon, 2010; Garcia-Argibay et al., 2019; Ingendoh et al., 2023; Lee-Harris et al., 2018; Orozco et al., 2020; Sharma et al., 2017). Heinrich Wilhelm Dove (1803–1879) discovered binaural beats in 1839 and Gerald Oster (1918-1993) tested them on the encephalograph (Filimon, 2010). The brain releases electric impulses, brain waves, which determine states of conscience. These waves are classified into four fundamental types: delta, theta, alpha, and beta, but the gamma frequency is also mentioned in the literature (2010).

It is claimed that the brain responds to binaural beat stimuli by synchronizing its own electrical cycles with the audio (Sharma et al., 2017). Although debated, the brainwave entrainment hypothesis states that external stimuli at specific frequencies cause the brain's electrocortical activity to oscillate at the same frequencies. These brain responses can be measured primarily using two strategies: auditory steady-state responses or auditory frequency-following responses (Ingendoh et al., 2023; Orozco et al., 2020). As such, "the cognitive effects of binaural beats are attributed to their capacity to drive neural oscillations at the beat frequency through differential hemispheric synchronization frequencies" (Orozco et al., 2020, p. 2). However, the evidence suggests that the effects of binaural beats are influenced by factors such as the frequency, the duration of exposure, and the type of masking used for binaural beats (Garcia-Argibay et al., 2019; Orozco et al., 2020).

Having said this, it is expected that binaural beats facilitate moving from one state of consciousness to another. Nowadays, individuals are predominantly in the beta wave frequency, which is useful for analytical thinking because it is an alert state, but also it implies tension, anxiety, and stress. Consequently, the brain usually attempts to transform beta waves into alpha ones, which are associated with calm and a receptive mind, supporting mental focus and the learning process (Filimon, 2010; Sharma et al., 2017).

Based on a meta-analysis, Garcia-Argibay et al., (2019) found that binaural beats can influence mental states and cognition. There is evidence suggesting that exposure to binaural beats leads to psychophysiological changes, such as reduced anxiety and stress levels, and increased creativity. Moreover, binaural beats are associated with improvements in memory, attention, and pain perception (Garcia-Argibay et al., 2019; Orozco et al., 2020).

Similarly, the role of binaural music (music embedded with binaural beats) has been studied in various fields, such as music therapy to enhance people's health (Filimon, 2010; Rahman et al., 2021). The evidence suggests that binaural music decreases anxiety in surgical patients, such as those undergoing cataract surgery (Wiwatwongwana et al., 2016) and helps manage hypertension (Wichian et al., 2021). Bae et al. (2023), specifically, demonstrated its anxiety-reducing effects in surgical patients.

While binaural music, i.e., "binaurally recorded sounds, binaural beats, a slow tempo, and gradual changes" (Lee-Harris et al., 2018, p. 1), is often associated with relaxation, stress and anxiety reduction, and improved focus, there is no specific scientific research on its impact on academic achievement. However, we can assume potential effects based on related research.

Low levels of stress positively influence attention and cognitive performance. Consequently, students who experience lower stress levels may be better equipped to concentrate, study effectively, and perform well academically (Algailani et al., 2023; McBride & Greeson, 2023; Ye et al., 2019). Similarly, anxiety can interfere with academic performance, for example, negatively affecting attention, memory, and problem-solving abilities. Therefore, low levels of anxiety create conditions for effective learning and academic performance (Pekrun et al., 2002). Furthermore, listening to binaural music may also enhance mood and increase motivation (Lee-Harris et al., 2018), positively affecting academic engagement and perseverance in the face of challenges to academic achievement (Mega et al., 2014; Pekrun et al., 2002).

Note that, primarily, anecdotal evidence supports the positive effects of binaural music on focus and relaxation for studying. Consequently, rigorous scientific research is needed to establish clear causation and quantify the extent of these effects. As with any intervention, it is advisable to consider individual needs, preferences, and the context in which binaural music is used concerning academic pursuits. As research in this field continues to advance, the study of the relationship between binaural music and academic achievement will contribute to the ongoing improvement of educational practices. Importantly, individual preferences play a significant role, and not everyone may benefit from or enjoy binaural music (Rahman et al., 2021).

In this context, this article contributes to the literature by testing the impact of binaural music on the educational outcomes of university students, using their classes of international economics and microeconomics and building longitudinal data from a natural experiment in times of covid-19.

3. Method

Artificial intelligence was occasionally used in different sections of the article exclusively to improve writing and readability (OpenAI, 2023).

The covid-19 crisis represented a succession of new challenges at all levels of education, demanding pedagogical innovations (Keržič et al., 2021; Nieto-Escamez & Roldán-Tapia, 2021). Accordingly, different strategies to capture and maintain students' attention were explored, including the use of background music at the beginning of the lockdown (from March to December 2020). This decision was a result of the necessity to choose between noise or music due to the disturbances present in the household environment, such as crying babies or unexpected interruptions. Initially, instrumental classical music was employed to reduce stress based on existing research (Pelletier, 2004), and the positive effects observed in both lecturing (Dosseville et al., 2012) and mathematical problem-solving tasks (Pavlyugina et al., 2012), as well as the Mozart effect (Beauvais, 2015). However, due to the exigencies of the crisis, the exploration of strategies led to the discovery of binaural music (Filimon, 2010; Rahman et al., 2021; S. Sharma et al., 2017). This shift from classical to binaural music resulted in a positive experience, seemingly improving the average academic performance of students. However, this perception was based solely on anecdotal evidence, generating a pilot study to gather data. In addition, students did not object to the use of background music; instead, they supported it. Consequently, a decision was made to document and formalize this natural experiment systematically in the first semester of 2021 (from January to June 2021), allowing for data collection as described below.[1]

3.1. Participants

The sample includes 73 undergraduate students, who were enrolled in the courses of Microeconomics (two groups: micro A and micro B) and International Economics. Thus, the participants were in three different groups: Microeconomics A ($n = 27$; 15 females and 12 males; average age 20.5), Microeconomics B ($n = 25$; 9 females and 16 males; average age 20.6), and International Economics ($n = 21$; 3 females and 18 males; average age 21.4). The students of Microeconomics were first-year undergraduate business/management students, and the students of International Economics were third year. English was the language of instruction, and the classes started on January 18 and finished on May 20 in the year 2021 (with final exams at the

beginning of June), in accordance with the academic calendar of Anonymous University.

3.2. Material

3.2.1. Online class

Given the covid-19 restrictions, Moodle was used as the learning platform and Google Meet for online classes. The lessons of Microeconomics A were on Mondays and Wednesdays, starting at 11:00 am and finishing at 12:30 pm, with 1 hour and 15 minutes of lecture, combined with synchronized slides, and using the last 15 minutes to answer general questions (if necessary). Microeconomics B was on Fridays, starting at 9:15 am, with a break from 10:30 am to 11:00 am, finishing the class at 12:30 pm, giving the last 15 minutes to general questions. In the case of International Economics, the lessons were on Tuesdays and Thursdays, starting at 7:00 am and finishing at 8:45 am, with 1 hour and 30 minutes of lecture, combined with synchronized slides, and using the last 15 minutes to answer general questions. The students had already received their classes in a very similar context in the year 2020, so they were familiar with the procedure. In addition, note that the groups are small, and the lecture is not traditional because students actively participate in class.

The topics taught in microeconomics are basic, based on Parkin's textbook. The course is divided into three parts. The first part includes the chapters on supply and demand, elasticity, and consumer theory. The second part includes the chapters on the theory of the firm, output, and costs. In partial three, the course finishes with the chapters on market structures (Parkin, 2019). Krugman's textbook is used in the case of International Economics. The first part includes chapters on international trade theory, e.g., comparative advantage, Heckscher-Ohlin model, and scale economies. The second part includes chapters on trade policy and international finance. In partial three, the course finishes with chapters on open macroeconomic policy (Krugman et al., 2018).

3.2.2. Multiple-choice quiz

At the end of each lecture of the first and second parts of the courses, in five minutes, the students answered a multiple-choice quiz (MCQ) in Moodle, which was created from the lecture material. The students answered the same questions, randomly assigned, and backtracking was prohibited to prevent copying. The students were informed that the results of the MCQs constituted 30% of their grades in the mid-term exams (partials one and two). Actually, the quizzes were only 10% of the final grade according to the syllabus, as many other assignments and partial exams were accounted for the final grade. However, this 10% figure was not emphasized to the students, who were expected to try their best on each quiz because of the 30% weighting.

In the case of Microeconomics, each MCQ consisted of three questions, two

of them type true/false, and one of them type multiple options (four answers, of which one was correct), in a 0-100 grading scale where the multiple option question equalled 40 points. In the case of International Economics, the three questions were multiple options (five answers, of which one was correct), in a 0-100 grading scale, where the value of the most difficult question (in opinion of experts) was 40 points. In total, there were 11 MCQs in the case of Microeconomics (six in the partial one, and five in the partial two) and 16 MCQs in the case of International Economics (eight in each partial).

The typical recommendations found in methodological textbooks to write the questions were followed: avoid vagueness, avoid excessively complicated questions, avoid negative stems, and use correct grammar. Nevertheless, the question types of the MCQs of Microeconomics and International Economics were different, as a response to students' background on economics and the requirements of their study programs. The students of Microeconomics passed an introductory course to economics and history of economic thought. In contrast, the students of International Economics additionally passed the courses of microeconomics and macroeconomics.

Note that partial three of the courses had a completely different pedagogical dynamic. The students were tasked with presenting topics and scientific articles, supplemented by the use of multimedia resources including videos, memes, and Kahoot quizzes. Additionally, the final exam for these courses consisted of approximately 80% of questions related to topics covered in the third partial, with the remaining 20% focused on material from the first and second partials.

3.2.3. Music

Binaural beats present two neighbouring frequencies to each ear, separately. For its part, binaural music, audio recorded using binaural beats technique, already presents the resulting frequency, summing each tone before the presentation. As a result, binaural music becomes monaural and is presented to both ears at the same time (Orozco et al., 2020). As such, this type of music was used as background music, played at low volumes to ensure it would not disturb the lecture. Importantly, Garcia-Argibay et al. (2019) show meta-regression results suggesting that it does not seem to be necessary to mask binaural beats with white noise or pink noise in terms of effectiveness. Therefore, the choice of masking method (white noise, pink noise, music) may not significantly impact the effectiveness of binaural beats, providing flexibility in how binaural beats are presented in various settings.

Thus, binaural music represents the treatment condition compared to the control condition of no music. Although students were not formally informed about the presence of music (though

they obviously noticed it), they listened to the music a few minutes before the start of class and during the lecture. Among the various options available on YouTube, a piece of music from Greenred Productions, with alpha binaural beat frequency, was selected after brainstorming, considering factors such as the number of views, likes, and comments.[2]

3.2.4. Previous tests on economics

At the beginning of the semester, all students completed general exams on economics to control for prior knowledge of the subject. In the case of Microeconomics, students answered in 75 minutes a test with 18 questions on the basic topics of chapters 1 and 2 of Parkin´s textbook. They answered the same questions, randomly assigned, and backtracking was prohibited to prevent copying, in a 0-100 grading scale (Microeconomics A: mean *(ME)* = 75.19, standard deviation *(SD)* = 13.41; Microeconomics B: ME = 82.71, SD = 16.83).

Similarly, in the case of International Economics, students answered three tests. Two of them on microeconomics, the first one on basic topics, the same test for students in Microeconomics (ME = 52.63, SD = 8.94), and the second one on supply, demand, and elasticity (20 questions, ME = 57.63, SD = 16.17). The third test was on macroeconomics: aggregate supply, aggregate demand, IS-LM model, and fiscal and monetary policies (25 questions, ME = 48.42, SD = 15.13).

3.2.5. Background questionnaire

After the mid-term exams and before the topics of the third partial began, all students completed a background questionnaire to provide control variables (see Table 1). In addition to general background data, they provided information and opinions (10-point rating scales) about the background music, difficulty of MCQs, difficulties because of covid, technical or internet issues, language concerns, and grade point average (GPA). Only at this moment were the students informed about the use of music and plans to use their data for research purposes (exclusively and on condition of anonymity). Accordingly, students provided written consent to use their MCQ results and background data. One student refused and was removed from the analysis.

3.3. Procedure

In the case of Microeconomics, binaural music was used during the lectures for group A in the first part of the course, making group B the control group. Conversely, binaural music was used during the lectures for group B in the second part of the course, with group A becoming the control group. In that manner, both groups operated as treatment and control groups, allowing intra-and inter-group comparisons. In the case of International Economics, binaural music was used in lectures during the first partial, with the quiz scores from the second partial serving as the data for the control group, allowing for only intertemporal comparisons (see Table 2).

TABLE 1. Background questionnaire.

		Micro A		Micro B		International Economics	
		ME	SD	ME	SD	ME	SD
Q1	Did you notice if background music was played during the sessions before the first midterm? (1 = Yes / 0 = No)	0.93	0.26	1.00	0.00	0.95	0.21
Q2	If you noticed, on a scale from 0 to 10, where 0 represents "not at all" and 10 represents "a lot", to what extent did this music bother you? Please write a number (if you didn't notice it, write -1).	1.63	2.15	1.40	2.12	2.95	3.39
Q3	If you noticed, do you have any comments about the mentioned music?	N/A	N/A	N/A	N/A	N/A	N/A
Q4	Did you notice that during the sessions before the second midterm exam, your professor did not play the mentioned background music? 1 = Yes / 0 = No.	0.67	0.47	N/A	N/A	0.67	0.47
Q5	On a scale from 0 to 10, where 0 is "very easy" and 10 is "very difficult", how difficult were the quizzes that were answered at the end of each lesson?	6.00	1.59	6.24	1.27	7.67	1.25
Q6	On a scale from 0 to 10, where 0 is "very easy" and 10 is "very difficult", do you think it was possible to cheat by answering the quiz, specifically, having someone pass the questions or answers to you?	5.67	2.80	7.84	1.46	7.95	1.99
Q7	On a scale from 0 to 10, where 0 is "strongly disagree" and 10 "strongly agree", do you think it was necessary to pay attention in class to be able to answer the quiz correctly?	9.41	0.91	9.36	1.05	9.19	1.53

Q8	On a scale from 0 to 10, where 0 is "very easy" and 10 is "very difficult", how difficult has it been for you to study the subject (specifically the economics class) as a result of the covid-19 crisis?	7.37	1.57	7.76	1.50	7.14	1.49
Q9	On a scale from 0 to 10, where 0 is "very easy" and 10 is "very difficult", how difficult has it been for you to study the subject due to problems with the internet connection, audio, microphone or other technical problem?	5.89	2.56	6.40	1.88	3.95	2.94
Q10	What device did you normally use to attend the online class? Computer / laptop = 1, Phone = 2, Tablet = 3, Other = 4.	1.13	0.42	1.22	0.57	1.21	0.40
Q11	On a scale from 0 to 10, where 0 is "very easy" and 10 is "very difficult", how difficult has it been for you to study the subject because the language of instruction is English?	4.52	2.85	3.60	3.09	4.00	3.48
Q12	Is English your mother tongue? 1 = Yes / 0 = No	0.07	0.26	0.08	0.27	0.00	0.00
Q13	On a scale from 0 to 10, where 0 is "very low" and 10 is "very high", how good do you consider your level of English, especially technical, to understand economics?	7.85	1.18	7.68	1.59	7.95	1.62
Q14	What was your GPA last semester?	8.79	0.41	8.34	0.70	8.43	0.45
Q15	So far, what is your GPA this semester?	8.68	0.46	8.19	0.76	8.19	0.63

Notes: *ME*: mean; *SD*: standard deviation. N/A: not applicable. Students in Microeconomics B listened to the music in the second part of the course. GPA: Grade Point Average (scale from 0 to 10).

TABLE 2. Design of the performed comparisons.

	Partial 1	Partial 2	
Microeconomics A	Experimental condition	Control condition	Intra-group comparison
Microeconomics B	Control Condition	Experimental condition	Intra-group comparison
Microeconomics A versus B	Inter-group comparison	Inter-group comparison	
International Economics	Experimental condition	Control condition	Intra-group comparison

Evidently, the students did not attend exactly the same lecture. While every effort was made to explain each topic in the same manner by the same instructor, and to avoid providing extra information that could assist one group over the other in answering the MCQs, each lecture differed due to student participation. This is a limitation of this study, yet the strategy is under real conditions. Moreover, consider the use of panel data, which offers greater degrees of freedom, increased variability, reduced collinearity, enhanced efficiency, and the most important aspect: control over the researcher effect by controlling for time-invariant variables.

3.4. Data analysis

To test the impact of binaural music on the MCQ scores, first, statistical descriptive analysis and t-tests to compare means between the experimental and the control groups (including the intertemporal observations, for intra-group analysis) were used. Second, regression models with a dummy variable coded 1 were considered for the treatment, while the initial exams on economics and several other variables from the background questionnaire operated as control variables. The correlation coefficients amongst independent variables were low (see Table 3). Nevertheless, not all possible independent variables were entered at the same time in the regression models because of multicollinearity concerns.

Accordingly, five baseline regression models are given by equations (1) to (5), which were estimated using random effects to allow the inclusion of time-invariant variables (also supported by the Hausman test).

MCQ represents the score of the quizzes by subject according to the subscript. The subscript First partial Microeconomics A vs B signifies the inter-group comparison where Microeconomics A received the *treatment* (listening to binaural music) and Microe-

TABLE 3. Correlation matrix.

	MCQ score	Treatment	Q1	Q2	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Micro 1	Micro 2	Macro
MCQ score (quiz)	1																		
Treatment	0.16	1																	
Q1	-0.06	-0.01	1																
Q2	-0.06	0.002	0.13	1															
Q4	-0.001	0.00	0.19	0.12	1														
Q5	-0.23	-0.004	0.12	0.32	-0.15	1													
Q6	-0.18	-0.03	0.05	-0.07	-0.02	0.30	1												
Q7	-0.02	0.01	0.02	-0.10	0.11	0.10	0.38	1											
Q8	-0.02	-0.005	0.18	0.27	-0.05	0.24	0.19	0.18	1										
Q9	0.15	-0.002	0.01	0.004	-0.22	-0.13	-0.17	-0.22	0.12	1									
Q10	0.01	-0.01	0.05	0.01	-0.06	-0.15	-0.04	-0.16	0.01	0.19	1								
Q11	-0.07	0.01	0.07	0.31	0.13	0.07	-0.06	0.02	0.17	0.33	0.03	1							
Q12	0.12	0.005	0.02	-0.06	-0.07	-0.08	-0.12	-0.10	-0.17	0.13	-0.09	-0.12	1						
Q13	0.06	0.001	-0.01	0.30	-0.20	0.17	0.00	-0.16	0.09	-0.17	-0.02	-0.45	0.16	1					
Q14	0.14	0.02	0.04	0.16	-0.11	-0.15	-0.14	-0.09	0.18	0.12	-0.16	-0.08	0.00	0.29	1				
Q15	0.18	0.02	-0.04	-0.001	-0.06	-0.22	-0.13	0.10	0.07	0.20	-0.12	-0.11	0.11	0.14	0.56	1			
Test on Micro 1	0.29	-0.01	-0.04	-0.10	-0.18	-0.32	-0.03	-0.001	0.04	0.46	-0.05	0.11	0.19	0.00	0.11	0.13	1		
Test on Micro 2	-0.06	0.00		-0.45	-0.01	-0.25	0.29	0.20	-0.07	-0.52	0.05	-0.27		-0.12	-0.07	-0.27	-0.04	1	
Test on Macro	0.09	0.00		-0.30	-0.01	0.08	0.13	0.07	0.11	-0.55	-0.07	-0.57		0.47	0.10	-0.07	-0.16	0.42	1

Note: See Table 1 for the definition of variables Q1 to Q15.

conomics B was the control group. The subscript Second partial Microeconomics A vs B signifies the inter-group comparison where Microeconomics A was the control group and Microeconomics B received the treatment. Subscript i ($i = 1, ..., n$) and subscript t ($t = 1, ..., t$) indicate the cross-section and time dimension. X represents control variables, β signifies the coefficients to be estimated, and e and v are the error terms.

$$MCQ_{it \text{ Score Microeconomics A}} = \beta_0 + \beta_1 Treatment_{it} + X'\beta_n + e_j + v_{jt} \qquad (1)$$

$$MCQ_{it \text{ Score Microeconomics B}} = \beta_0 + \beta_1 Treatment_{it} + X'\beta_n + e_j + v_{jt} \qquad (2)$$

$$MCQ_{it \text{ Score first partial Microeconomics A vs B}} \qquad (3)$$

$$= \beta_0 + \beta_1 Treatment_{it} + X'\beta_n + e_j + v_{jt}$$

$$MCQ_{it \text{ Score second partial Microeconomics A vs B}} \qquad (4)$$

$$= \beta_0 + \beta_1 Treatment_{it} + X'\beta_n + e_i + v_{it}$$

$$MCQ_{it \text{ Score International Economics}} = \beta_0 + \beta_1 Treatment_{it} + X'\beta_n + e_j + v_{jt} \qquad (5)$$

4. Results

Descriptive statistics and t-tests for comparing two means are shown in Table 4. In the case of intra-group comparisons between Microeconomics A and B, the t-tests indicate statistically different means, albeit with contradictory results. Specifically, for Microeconomics A, the evidence suggests higher MCQ scores for the control group, that is, higher scores for the second partial without music. Conversely, for Microeconomics B, the evidence suggests higher MCQ scores for the treatment group, that is, higher scores for the second partial with music.

In the case of inter-group, comparisons between Microeconomics A and B, for the first partial, the mean comparison between Microeconomics A ($ME = 89.16$) and Microeconomics B ($ME = 82.17$) suggests significantly higher scores for the treatment group (Microeconomics A) than for the control group (Microeconomics B). However, in the second partial, the mean comparison between Microeconomics A ($ME = 94.10$) and Microeconomics B ($ME = 90.23$) indicates statistically higher scores for the control group (Microeconomics A), with a significance level of 8% (thus, an insignificant difference under stricter p-values).

In contrast, in the case of International Economics (an intra-group comparison), a statistically significant difference in the means is evident, favouring the treatment. The mean score in the second partial was 58.05, while the mean score in the first partial (with music) was 76.67.

TABLE 4. Descriptive statistics and two-sample t test (with equal variances).

Microeconomía A

Group	Obs.	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Control (0)	117	94.10	1.26	13.59	91.61	96.59
Treatment (1)	143	89.16	1.42	16.93	86.36	91.96
Combined	260	91.38	0.97	15.68	89.47	93.30
Difference		4.94	1.93		1.13	8.75

Difference = mean (0) - mean (1)   Degrees of freedom = 258
Ho = 0   Ha ≠ 0 : Pr(|T| > |t|) = 0.01
t = 2.55

Microeconomics B

Group	Obs.	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Control (0)	106	82.17	2.55	26.22	77.12	87.22
Treatment (1)	87	90.23	1.93	17.98	86.40	94.06
Combined	193	85.80	1.67	23.17	82.51	89.09
Difference		-8.06	3.31		-14.59	-1.53

Difference = mean (0) - mean (1)   Degrees of freedom = 191
Ho = 0   Ha ≠ 0 : Pr(|T| > |t|) = 0.02
t = -2.43

Microeconomics first partial A vs B (group B control)

Group	Obs.	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Control (0)	106	82.17	2.55	26.22	77.12	87.22
Treatment (1)	143	89.16	1.42	16.93	86.36	91.96
Combined	249	86.18	1.37	21.62	83.49	88.88
Difference		-6.99	2.74		-12.39	-1.59

Difference = mean (0) - mean (1)   Degrees of freedom = 247
Ho = 0   Ha ≠ 0 : Pr(|T| > |t|) = 0.01
t = -2.55

Microeconomics second partial A vs B (group A control)

Group	Obs.	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Control (0)	117	94.10	1.26	13.59	91.61	96.59
Treatment (1)	87	90.23	1.93	17.98	86.40	94.06
Combined	204	92.45	1.10	15.69	90.28	94.62
Difference		3.87	2.21		-0.48	8.23

Difference = mean(0) - mean(1)   Degrees of freedom = 202
Ho = 0   Ha ≠ 0 : Pr(|T| > |t|) = 0.08
t = 1.75

International Economics

Group	Obs.	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Control (0)	154	58.05	2.33	28.97	53.44	62.66
Treatment (1)	159	76.67	2.17	27.32	72.39	80.95
Combined	313	67.51	1.67	29.61	64.22	70.80
Difference		-18.61	3.18		-24.88	-12.35

Difference = mean (0) - mean (1)   Degrees of freedom = 311
Ho = 0   Ha ≠ 0 : Pr(|T| > |t|) = 0.00
t = -5.85

Table 5 shows the main regression results of the baseline models. Overall, the regression results align with the findings of the t-tests, indicating that the evidence does not support the effectiveness of the treatment in the case of Microeconomics A, as observed in the intra-group analysis (column 1). It appears that the music had a negative effect on the MCQ scores, decreasing the average score by about five points. In contrast, the treatment shows a positive impact, resulting in an average increase of eight points in the case of Microeconomics B, as observed in the intra-group analysis (column 2).

In the first partial (column 3), where Microeconomics A serves as the treatment group and Microeconomics B as the control group, the inter-group analysis indicates a positive impact of the treatment, amounting to approximately eight points. Conversely, in the second partial (column 4), where Microeconomics B is designated as the treatment group and Microeconomics A as the control group, the coefficient of Treatment lacks statistical significance.

In the case of International Economics (column 5), the regression results suggest a positive impact of the treatment, approximately 20 points. It is also noteworthy that the majority of control variables are statistically significant only in the case of International Economics. In other words, most of the control variables lack relevance in explaining MCQ scores. This makes sense if we consider that it is the lecture content that should primarily explain quiz results.

As additional robustness checks, the analysis was replicated using fixed effects (without control variables that are time–invariant), and the coefficients of the treatment were qualitatively the same. Moreover, the analysis was replicated including different combinations of control variables (using the questions in Table 1) and time dummies. The use of time dummies negatively affected the significance of the treatment, but it was clear that time dummies overlapped the treatment presenting statistical significance and the expected sign in the treatment times (excepting Microeconomics A). In addition, in the case of Microeconomics, the analysis was replicated using gender subsamples, and the findings are qualitatively the same (not reported in tables).

Due to the nature of MCQ scores, which might tend to provide discrete variables (especially in the case of Microeconomics due to true/false questions), the regression analysis was replicated using logit regressions. MCQs with a score of 100 were coded as 1, and 0 otherwise. These results are reported in Table 6. Overall, the major findings are similar. Particularly, in the case of Microeconomics, the evidence is mixed.

TABLE 5. Regression results.

	(1) Microeconomics A	(2) Microeconomics B	(3) First partial Microeconomics A vs. B	(4) Second partial Microeconomics A vs. B	(5) International Economics
Treatment	-5.01***	8.32**	8.12***	-4.03	20.48***
Q2 (music disturbs)	0.18	1.05	0.31	-0.08	1.00
Q7 (pay attention)	0.02	-2.55	-1.00	-0.11	-0.30
Q8 (covid-19 issues)	-0.18	1.53	-0.57	0.62	-4.08***
Q13 (English concerns)	0.54	0.90	1.57	0.40	-0.89
Q15 (GPA current semester)	-1.08	5.13	0.21	1.97	7.88***
Test on micro 1	0.03	0.03	0.03	0.004	0.20
Test on micro 2					-0.14
Test on macro					0.42**
Constant	98.04***	39.69	76.99***	70.24***	5.04
Observations	249	172	232	189	286
$n \times t$	23 x 11	17 x 11	40 x 6	39 x 5	19 x 16
r^2	0.03	0.06	0.05	0.03	0.19

Note: random-effects GLS regression. (*) [**] and {***} indicate statistical significance at the (10%) [5%] and {1%} levels.

TABLE 6. Logit regression results.

	(1) Microeconomics A	(2) Microeconomics B	(3) First partial Microeconomics A vs. B	(4) Second partial Microeconomics A vs. B	(5) Economía Internacional
Treatment	-0.85***	0.52$^\psi$	0.15	-0.56	1.67***
Q2 (MUSIC disturbs)	0.05	0.08	0.07	-0.02	-0.04
Q7 (PAY attention)	-0.02	-0.26	-0.01	-0.16	0.07
Q8 (COVID-19 issues)	-0.07	0.13	-0.09	0.10	-0.23*
Q13 (English concerns)	0.10	0.09	0.16	0.05	0.05
Q15 (GPA current semester)	-0.17	0.39	-0.06	0.24	0.40*
Test on micro 1	0.002	0.0003	-0.002	0.003	0.02
Test on micro 2					-0.005
Test on macro					0.01
Constant	2.72	-2.07	0.60	-0.24	-5.51**
Observations	249	172	232	189	286
$n \times t$	23 x 11	17 x 11	40 x 6	39 x 5	19 x 16

Notes: logit regressions report coefficients. The odds ratio is calculated as exp(β), where β represents the regression coefficient. (*) [**] and {***} indicate statistical significance at the (10%) [5%] and {1%} levels. $^\psi$ p-value = 0.126

5. Discussion

The results are robust in the case of International Economics. In contrast, the evidence is somewhat mixed in the case of the groups in Microeconomics. It is possible that the type of MCQs used to measure academic performance in these classes were not the most appropriate, particularly due to the inclusion of true/false questions where students have a 50% chance of guessing the correct answer. Regarding this issue, there is little else to be done, as the quizzes are prepared with the study program in mind rather than for experimental design purposes to test the impact of binaural music. It is important to remember that the data for this research come from a natural experiment.

In light of the argument, the background questionnaire provides more information to explain the concerns in the case of Microeconomics, and in favour of binaural music. There is an open question: "do you have any comments on the mentioned music?" In general, students enjoyed the music, typically commenting "relaxing", "it relaxed me", "I like it", "I liked it, it attracted my attention in class", "I thought it was a good idea, personally it didn't bother me or anything, on the contrary it was comfortable", "It is a very calm music that after listening to it for a while I forget about it". These comments agree with previous empirical findings suggesting that binaural music contributes to relaxation, particularly, in the case of young individuals (Lee-Harris et al., 2018).

However, a few students mentioned some kind of discomfort (one student from International Economics, four students from Microeconomics A and two students from Microeconomics B). In the case of microeconomics students, many times, the negative comment on the music was related to the quality of internet connection and sound, for example "It was distorted by the microphone and it was heard badly" "A little annoying because of the platform, it would be better if each student plays such music on their own" "Sometimes the volume of the music was very loud, I would have liked another type of music (more instrumental) but that is already an individual case" "it was heard badly and from time to time if it came to bother". Therefore, in Microeconomics, the mixed results of the effects of binaural music on academic performance could be related to the quality of internet connection and sound.

Overall, the open questions of the background questionnaire, and the statistical and regression results support the expectation of a positive influence of binaural music on academic performance. Since there are no other studies with a similar goal, it is not possible to compare and discuss the current results with prior literature specific to this kind of music. Nonetheless, in general, the results align with the Mozart effect and its sociocultural implications, while care must be taken not to interpret the findings as magical powers (Beauvais, 2015). Additionally, the findings are consistent with studies

suggesting a positive correlation between music, concentration, relaxation, mood, and educational outcomes (Antony et al., 2018; Dosseville et al., 2012; Ishiguro et al., 2023; Pavlyugina et al., 2012). In particular, the results agree with the role of instrumental classical music in lectures, which enhances the learning environment and leads to improved quiz scores, ultimately contributing to higher educational attainment (Dosseville et al., 2012).

As such, binaural music helps to focus and improves academic performance. Moreover, its positive effects are long-term, spanning practically over one semester. In contrast, previous studies have utilized classical music and data from a single lecture or exercise, providing results at a single point in time and on a short-term basis (Dosseville et al., 2012; Pavlyugina et al., 2012).

Nevertheless, it is clear that not all students enjoy binaural music (given the general comments of students cited above), which agrees with previous studies (Rahman et al., 2021). Multiple factors may be affecting the emotions and responses of listeners to binaural music; the literature has emphasized familiarity and preferences (Lee-Harris et al., 2018). Furthermore, the evidence suggests that not all individuals respond in the exact same way to the same binaural music, because individuals have different states of brainwaves (Sharma et al., 2017). Therefore, even when the policy suggestion could be to

use the music in the traditional teaching-learning process, in the classroom, hybrid or not, it is more important that students are aware of the positive impact of binaural music to focus, and so to study. Thus, as some students commented in the background questionnaire, individually, each student should play the music while studying, according to their personal preferences.

5.1. Key limitations of this research and further research

First, while panel data offer numerous advantages and both dependent and independent variables are directly measured, certain control variables included in the analysis, such as grade point average, internet issues, and language concerns, are observational and self-reported. Second, explanatory variables that are time-invariant or change slowly, such as psychological variables, socioeconomic status (parental education, occupation, and income), cultural capital, and social capital, are not directly observed in the dataset. Instead, they are controlled using panel data and a quasi-experimental design. However, it would be desirable to include these variables in future research. Moreover, due to the online nature of the classes, it was not possible to ensure that students were both attending the lecture and listening to the music. Consequently, causality remains a potential concern. Therefore, further research employing a larger sample size and strict experimental designs while carefully addressing ethical concerns is necessary to better control for

causality. Nevertheless, supporting this research, it is noteworthy that general student comments about the music align with the statistical findings. The majority of students acknowledged that binaural music is relaxing and helps them to focus.

Future research should also assess the role of other musical genres, their associations with different subjects, and their effects over different timeframes. Consequently, further research should analyse the relationship between music preferences and academic performance. Moreover, future studies could provide additional insights into the benefits of binaural music by examining its effects in various contexts. For example, testing the impact of binaural music in the workplace, particularly in jobs requiring high levels of concentration. Accordingly, it would be interesting to develop other types of intervention studies using binaural music to support the major findings of this research and contribute to the development of public policies. Borrowing other words, "music is present in the everyday life of humans, and it can be used to increase the efficiency of many kinds of professional activity" (Pavlyugina et al., 2012: 354).

6. Final remarks

The classroom has been changing over the last 20 years, incorporating new tools to aid in the learning-teaching process, such as projectors, computers, and the internet. Generally, the learning-teaching process evolves over time, but the pandemic has accelerated unforeseen changes, particularly in technological advancements (Keržič et al., 2021; Nieto-Escamez & Roldán-Tapia, 2021). In this context, the findings of this research can be added to the extensive list of new and positive didactic strategies. The evidence suggests that binaural music can be highly useful as a didactic tool in the classroom, whether virtual or not, or as an additional tactic for students to focus and study at home. Accordingly, the findings of this research recommend the use of binaural background music in the classroom, analogous to music used during physical exercise or shopping at the mall. However, instructors should regulate its use based on students' signals or gestures indicating discomfort. This strategy can help students focus on the lecture, which is beneficial for some students today who appreciate frequent breaks or shorter attention spans (Bradbury, 2016).

Notes

[1] The data set for replication and direct outcomes from the software are available at the following link https://figshare.com/s/73b74b06ecb488cba7d4

[2] This piece is available at the following link: https://www.youtube.com/watch?v=mg-7netw1JuM&t=3439s

[3] Only at this moment, in the Microeconomics groups, it was noted that a few students systematically answered the MCQs quickly (in some cases after a few seconds). Therefore, these observations were removed from the analysis (outliers). By contrast, in the International Economics class, students even asked for one or two more minutes to answer the MCQs.

Author's contributions

Edgar Demetrio Tovar-García: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project Administration; Resources; Software; Validation; Visualization; Writing (original draft); Writing (review & editing).

Acknowledgements

Competing interests

The author declares no competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in this manuscript. The music selection used in this study was composed and recorded by Greenred Productions, but is freely available to the public. As such, Greenred Productions has no financial conflict of interest.

Ethical approval

Ethical approval was not sought due to issues with access to ethical review boards during the Covid-19 pandemic. Given the crisis and the online teaching processes, this approach aligns with the requirements of the institution where the research was conducted and with general didactic tools supported by the institution. Furthermore, the study primarily involved a non-invasive intervention of background music, which participants willingly engaged with, as evidenced by their support for its implementation (and written consent). No adverse effects or objections were reported by the students throughout the study. It is worth noting the absence of any experimental conditions that could potentially pose risks to the well-being of participants, emphasizing the notion of a safe and ethically conducted study. Overall, the positive response and lack of negative outcomes from both participants and the research process itself warrant the conclusion that no significant harms were associated with participation in this study and subsequent publication. Therefore, no significant harms were associated with participation in this research. It is important to note that the collected data are used at an aggregate level and anonymously.

Informed consent

All participants provided written informed consent.

Data available at: https://figshare.com/s/73b74b06ecb488cba7d4

References

Albar, S. B., & Southcott, J. E. (2021). Problem and project-based learning through an investigation lesson: Significant gains in creative thinking behaviour within the Australian foundation (preparatory) classroom. *Thinking Skills and Creativity, 41*, 100853. https://doi.org/10.1016/j.tsc.2021.100853

Algailani, U. F., Tigabu, B. M., Rahim, Y. R., Alzbeede, A. A., & Alshaikhli, L. O. (2023). The impact of music on stress reduction and academic performance of dental students. *Cureus, 15* (10), e46554. https://doi.org/10.7759/cureus.46554

Antony, M., Priya, V. V., & Gayathri, R. (2018). Effect of music on academic performance of college students. *Drug Invention Today, 10* (10), 2093-2096.

Bae, J., Yoo, S., Kim, H., Kim, Y., Kim, J.-T., Lim, Y.-J., & Kim, H.-S. (2023). Effect of real-time binaural music on sedation with dexmedetomidine during spinal anesthesia: A triple-arm, assessor-blind, randomized controlled trial. *Journal of Clinical Anesthesia, 84*, 110997. https://doi.org/10.1016/j.jclinane.2022.110997

Baker, D., Hallam, S., & Rogers, K. (2023). Does learning to play an instrument have an impact on change in attainment from age 11 to 16? *British Journal of Music Education, 40* (3), 297-310. https://doi.org/10.1017/S0265051723000116

Beauvais, C. (2015). The 'Mozart effect': A sociological reappraisal. *Cultural Sociology, 9* (2), 185-202. https://doi.org/10.1177/1749975514557096

Blasco-Magraner, J. S., Marín-Liébana, P., & Botella-Nicolás, A. M. (2022). Efectos de la educación musical en el desarrollo emocional de adolescentes entre 10 y 18 años. Una revisión sistemática [Effects of the educational use of music on 10-to-18-year-old adolescents' emotional development. A systematic review]. *Música Hodie, 22*, e68847. https://doi.org/10.5216/mh.v22.68847

Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education, 40* (4), 509-513. https://doi.org/10.1152/advan.00109.2016

Brazhnikova, Y. A. (2016). Innovative technique of working with music text by symmetry operations. *Integration of Education, 20* (4), 507-521. https://doi.org/10.15507/1991-9468.085.020.201604.507-521

Briscoe, J. R., Daugherty, C. D., Terui, S., & Goldsmith, J. V. (2022). Theory in praxis: Using the modified version of the Social Network Strategy (mSNS) as a pedagogical tool for understanding. *Communication Teacher, 36* (4), 281-288. https://doi.org/10.1080/17404622.2022.2035420

Burnard, P., & Younker, B. A. (2004). Problem-solving and creativity: Insights from students' individual composing pathways. *International Journal of Music Education, 22* (1), 59-76. https://doi.org/10.1177/0255761404042375

Castillo-Montoya, M. (2019). Professors' pedagogical strategies for teaching through diversity. *The Review of Higher Education, 42* (5), 199-226. https://doi.org/10.1353/rhe.2019.0050

Cheah, Y., Wong, H. K., Spitzer, M., & Coutinho, E. (2022). Background music and cognitive task performance: A systematic review of task, music, and population impact. *Music y Science, 5*, 1-44. https://doi.org/10.1177/20592043221134392

Dosseville, F., Laborde, S., & Scelles, N. (2012). Music during lectures: Will students learn better? *Learning and Individual Differences, 22* (2), 258-262. https://doi.org/10.1016/j.lindif.2011.10.004

Filimon, R. C. (2010). Beneficial subliminal music: Binaural beats, hemi-sync and metamusic. In V. Munteanu, R. Raducanu, G. Dutica, A, Croitoru, & V. E. Balas (Eds.), *Recent Advances in Acoustics & Music. Proceedings of the 11th WSEAS International Conference on Acoustics and Music: Theory and Applications (AMTA '10)* (pp. 103-108). WSEAS Press. http://wseas.us/e-library/conferences/2010/Iasi/AMTA/AMTA-18.pdf

Garcia-Argibay, M., Santed, M. A., & Reales, J. M. (2019). Efficacy of binaural auditory beats in cognition, anxiety, and pain perception: A meta-analysis. *Psychological Research, 83* (2), 357-372. https://doi.org/10.1007/s00426-018-1066-8

Gershon, W. S., & Ben-Horin, O. (2014). Deepening inquiry: What processes of making music can teach us about creativity and ontology for inquiry based science education. *International Journal of Education y the Arts, 15* (19), 38.

Gillanders, C., & Casal De La Fuente, L. (2020). Enhancing mathematical thinking in early childhood through music. *Pedagogies, 15* (1), 60-79. https://doi.org/10.1080/155448 0X.2019.1673167

Guhn, M., Emerson, S. D., & Gouzouasis, P. (2019). A population-level analysis of associations between school music participation and academic achievement. *Journal of Educational Psychology, 112* (2), 308-328. https://doi.org/10.1037/edu0000376

Halawa, S., Hsu, Y.-S., Zhang, W.-X., Kuo, Y.-R., & Wu, J.-Y. (2020). Features and trends of teaching strategies for scientific practices from a review of 2008-2017 articles. *International Journal of Science Education, 42* (7), 1183-1206. https://doi.org/10.1080/09500693.2020.1752415

Hallam, S., & Price, J. (1998). Can the use of background music improve the behaviour and academic performance of children with emotional and behavioural difficulties? *British Journal of Special Education, 25* (2), 88-91. https://doi.org/10.1111/1467-8527.t01-1-00063

Hickey, P., Barnett-Young, A., Patel, A. D., & Race, E. (2020). Environmental rhythms orchestrate neural activity at multiple stages of processing during memory encoding: Evidence from event-related potentials. *PLoS ONE, 15* (11), e0234668. https://doi.org/10.1371/journal.pone.0234668

Holmes, S. (2021). The wider cognitive benefits of engagement with music. In A. Creech, D. A. Hodges, S. Hallam (Eds.), *Routledge international handbook of music psychology in education and the community*. Routledge.

Ilari, B. (2020). Longitudinal research on music education and child development: Contributions and challenges. *Music and Science, 3*, 1-21. https://doi.org/10.1177/2059204320937224

Ilari, B., Perez, P., Wood, A., & Habibi, A. (2019). The role of community-based music and sports programmes in parental views of children's social skills and personality. *International Journal of Community Music, 12* (1), 35-56. https://doi.org/10.1386/ijcm.12.1.35_1

Ingendoh, R. M., Posny, E. S., & Heine, A. (2023). Binaural beats to entrain the brain? A systematic review of the effects of binaural beat stimulation on brain oscillatory activity, and the implications for psychological research and intervention. *PLOS ONE, 18* (5), e0286023. https://doi.org/10.1371/journal.pone.0286023

Ishiguro, C., Ishihara, T., & Morita, N. (2023). Extracurricular music and visual arts activities are related to academic performance improvement in school-aged children. *Npj Science of Learning, 8* (1), 7. https://doi.org/10.1038/s41539-023-00155-0

Jaschke, A. C., Eggermont, L. H. P., Honing, H., & Scherder, E. J. A. (2013). Music education and its effect on intellectual abilities in children: A systematic review. *Reviews in the Neurosciences, 24* (6), 665-675. https://doi.org/10.1515/revneuro-2013-0023

Keržič, D., Alex, J. K., Pamela Balbontín Alvarado, R., Bezerra, D. da S., Cheraghi, M., Dobrowolska, B., Fagbamigbe, A. F., Faris, M. E., França, T., González-Fernández, B., Gonzalez-Robledo, L. M., Inasius, F., Kar, S. K., Lazányi, K., Lazăr, F., Machin-Mastromatteo, J. D., Marôco, J., Marques, B. P., Mejía-Rodríguez, O., … Aristovnik, A. (2021). Academic student satisfaction and perceived performance in the e-learning environment during the COVID-19 pandemic: Evidence across ten countries. *PLOS ONE, 16* (10), e0258807. https://doi.org/10.1371/journal.pone.0258807

Koponen, J. (2019). The flipped classroom approach for teaching cross-cultural communication to millennials. *Journal of Teaching in International Business, 30* (2), 102-124. https://doi.org/10.1080/08975930.2019.1663776

Krugman, P. R., Obstfeld, M., & Melitz, M. J. (2018). International economics: *Theory and policy* (11th ed.). Pearson.

Lee-Harris, G., Timmers, R., Humberstone, N., & Blackburn, D. (2018). Music for relaxation: A comparison across two age groups. *Journal of Music Therapy, 55* (4), 439-462. https://doi.org/10.1093/jmt/thy016

McBride, E. E., & Greeson, J. M. (2023). Mindfulness, cognitive functioning, and academic achievement in college students: The mediating role of stress. *Current Psychology, 42* (13), 10924-10934. https://doi.org/10.1007/s12144-021-02340-z

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106* (1), 121-131. https://doi.org/10.1037/a0033546

Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. *Psychological Science, 10* (4), 370-373. https://doi.org/10.1111/1467-9280.00170

Nieto-Escamez, F. A., & Roldán-Tapia, M. D. (2021). Gamification as online teaching strategy during covid-19: A mini-review. *Frontiers in Psychology, 12*, 648552. https://doi.org/10.3389/fpsyg.2021.648552

Orozco, H. D., Dumas, G., & Lehmann, A. (2020). Binaural beats through the auditory pathway: From brainstem to connectivity patterns. *ENeuro, 7* (2), 1-18. https://doi.org/10.1523/ENEURO.0232-19.2020

OpenAI. (2023). *ChatGPT (14 March version) [Comprehensive language model]*. https://chat.openai.com/chat

Parkin, M. (2019). *Microeconomics* (13th ed.). Pearson.

Pavlyugina, R. A., Karamysheva, N. N., Sakharov, D. S., & Davydov, V. I. (2012). Influence of music on the solution of mathematical logical tasks. *Human Physiology, 38* (4), 354-360. https://doi.org/10.1134/S0362119712030097

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37* (2), 91-105. https://doi.org/10.1207/S15326985EP3702_4

Pelletier, C. L. (2004). The effect of music on decreasing arousal due to stress: A meta-analysis. *Journal of Music Therapy, 41* (3), 192-214. https://doi.org/10.1093/jmt/41.3.192

Rahman, J. S., Gedeon, T., Caldwell, S., & Jones, R. (2021). Can binaural beats increase your focus? Exploring the effects of music in participants' conscious and brain activity responses. *Conference on Human Factors in Computing Systems - Proceedings, May*, 6. https://doi.org/10.1145/3411763.3451813

Rauscher, F., Shaw, G., & Ky, C. (1993). Music and spatial task performance. *Nature, 365*, 611. https://doi.org/10.1038/365611a0

Rose, D., Jones, A., & Heaton, P. (2019). Measuring the impact of musical learning on cognitive, behavioural and socio-emotional wellbeing development in children. *Psychology of Music, 47* (2), 284-303. https://doi.org/10.1177/0305735617744887

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y., & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44* (10), 1436-1460. https://doi.org/10.1002/tea.20212

Sharma, M., Biros, D., Ayyalasomayajula, S., & Dalal, N. (2020). Teaching programming to the post-millennial generation: Pedagogic considerations for an IS course. *Journal of Information Systems Education, 31* (2), 96-105. http://jise.org/Volume31/n2/JISEv31n2p96.html

Sharma, S., Rewadkar, S., Pawar, H., Deokar, V., & Lomte, V. M. (2017). Survey on binaural beats and background music for increased focus and relaxation. In *2017 International Conference on Emerging Trends and Innovation in ICT*, (pp. 98-103). IEEE. https://doi.org/10.1109/ETIICT.2017.7977018

Stough, C., Kerkin, B., Bates, T., & Mangan, G. (1994). Music and spatial IQ. *Personality and Individual Differences, 17* (5), 695. https://doi.org/10.1016/0191-8869(94)90145-7

Suárez, L., Elangovan, S., & Au, A. (2016). Cross-sectional study on the relationship between music training and working memory in adults. *Australian Journal of Psychology, 68* (1), 38-46. https://doi.org/10.1111/ajpy.12087

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science, 12* (3), 248-251. https://doi.org/10.1111/1467-9280.00345

Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist, 41* (4), 207-225. https://doi.org/10.1207/s15326985ep4104_1

Wichian, S. N., Klaphajone, J., & Phrompayak, D. (2021). Effects of music embedded with binaural and superimposed beats controlling hypertension in older adults: A quasi experimental study. *Pacific Rim International Journal of Nursing Research, 25* (3), 345-358.

Wiwatwongwana, D., Vichitvejpaisal, P., Thaikruea, L., Klaphajone, J., Tantong, A., & Wiwatwongwana, A. (2016). The effect of music with and without binaural beat audio on operative anxiety in patients undergoing cataract surgery: A randomized controlled trial. *Eye (Basingstoke), 30* (11), 1407-1414. https://doi.org/10.1038/eye.2016.160

Ye, L., Posada, A., & Liu, Y. (2019). A review on the relationship between chinese adolescents' stress and academic achievement. *New Directions for Child and Adolescent Development, 163*, 81-95. https://doi.org/10.1002/cad.20265

Author's biography

Edgar Demetrio Tovar-García. PhD in Economics from the University of Barcelona. He has worked as a Short-Term Consultant for the World Bank and held roles as a Research Professor at several universities in Mexico and Russia. He is a member of the National System of Researchers in Mexico (SNII level II). His research interests include the economics of education, language policy, the economic aspects of religion, and entrepreneurship, among others. He pursues a diverse, multidisciplinary research approach, emphasizing policy implications and econometric modelling, particularly regression analysis employing panel data.

 https://orcid.org/0000-0002-8826-0004

Book Reviews

Arthur, J. & Fernández, V. (2023).
En vista del fin excelente: la educación del carácter en las universidades [In view of the excellent goal: Character education in universities]
(Gonzalo Moreno-Ochoa)

Ahedo, J. & Arteaga-Martínez, B. (Eds.) (2023).
El tetraedro de la amistad. Estudiante, familia, escuela y sociedad [The tetrahedron of friendship. Student, family, school and society]
(Laura Guerrero-Puerta)

Martín-Gutiérrez, Á. (2020).
La escuela y su entorno en la sociedad del conocimiento. Un estudio focalizado en la Formación Profesional [The school and its environment in the knowledge society. A study focussed on Professional Training]
(Ángel-Freddy Rodríguez-Torres)

Book reviews

Arthur, J. & Fernández, V. (2023).
En vista del fin excelente: la educación del carácter en las universidades [In view of the excellent goal: Character education in universities].
Aula Magna Proyecto Clave McGraw Hill.
234 pp.

In this work, the crucial importance of universities in the formation of students is presented and questioned. The authors help us reflect on the central role of educators in shaping young people's character in a university environment and on how human flourishing can be achieved through the virtues of faith, hope, and charity. This work stands out for its deeply practical character, since as well as meditating on these topics, it suggests and describes specific measures that can provide great inspiration for higher-education centres.

The first chapter questions the influence of universities on students' flourishing. Universities must not only teach moral virtues but also how to put them into practice to ensure that students become people, citizens, professionals who contribute to the good of society. As the authors note, "there is a broad gulf between knowing and doing" (p. 29). Human beings need information, but to achieve a fulfilled life, we must put moral virtues into practice through understanding and prudence. The second of these, understood as practical wisdom (*phronesis*), plays a fundamental role, as it represents the capacity to act and decide what is best at each moment. Nonetheless, it is noted how universities, despite being aware that the education that they provide influences the character of young people, focus on knowledge and competences of an economic or professional nature. In other words, they have taken a certain market-based approach: the teachers offer knowledge (a product) while students seem to want nothing other than accreditation so that they can move on in their careers. The authors argue for the central role of universities in the formation of character, something that is sufficiently important that it should not be left to chance. Universities must embody values and encourage their students to propose a committed life for themselves. In particular, in a Catholic university, the

doctrinal foundation must inspire character education and guide what it does and what it teaches. A university is not Catholic because of its name nor because it teaches the doctrine of the Church; what makes it truly Catholic is the use of reason in all facets of university life, in how truth and love are pursued.

The second chapter considers character education in Spain. It is striking that the trend in English-speaking countries of incorporating training of the character, of moral virtues, into school and university educational projects has not taken root in Spain. One possible reason is the legislative framework. From the enactment of the LOECE (Organic Act governing the Status of Educational Centres) in 1980 to the present day with the LOMLOE (Organic Act amending the Organic Education Act), eight education acts and reforms have been passed. When discussing the new challenges facing Spanish universities, the authors rightly argue that ethics, values and virtues are the principle of progress. Therefore, a model that contributes to shaping people's character and identity must be proposed, so that our young people can "come *to be someone* and not just come to know how *to do something*" (p. 45).

In chapter three, they summarise the document *Character education in the universities: A framework for flourishing*, from the Jubilee Centre for Character and Virtues and the Oxford Character Project, the purpose of which is to show universities how to articulate and structure their mission to facilitate the flourishing of students and holistic development of charac-

ter. Universities, whether they set out to or not, shape the character of their students. A good education in advanced studies not only provides people with training in how to do or access a good job, but it must also influence what *they will become* and how they can contribute to society. Therefore, character education must permeate university life: culture, teaching, research, extracurricular activities, vocational guidance services, admissions, etc. In each of these points, the authors, aptly, propose ideas that invite us to reflect and to propose improvement activities. Along with this, specific measures are defined so that all of the staff of the university contribute to putting into practice this *framework for flourishing*.

Closely connected to the above, the need to teach professional ethics is discussed in chapter four. These should not be based on deontological codes or codes of conduct nor on the use of decision-making models. No: professional practice must be incorporated as an intrinsically moral whole including a concept of responsibility that inspires responses from everyone that are in accordance with what is good. The authors argue for replacing the concept of *professional career* with *professional vocation*, in which the person is involved as such in its exercise. It is also in work where the character of the person will be displayed, as *being* will be expressed through *doing*. Professional ethics are not a matter of avoiding evil. They are much more ambitious: they are a case of pursuing good.

Chapters five and six discuss the vocation of educating and of training teachers. The task of the teacher is fascinating. It is

defined as a service vocation that must start with transformation of oneself to become a virtuous educator. It requires serving others through committing oneself. The call to educate is not so much *doing things* as *living because of something* and *for someone*. The teacher, consequently, is not only a transmitter of content. Educators shape; through their mere presence, they influence the student, they cannot be neutral. Therefore, it is vital for them to have well-formed characters, as people cannot give what they do not have. So, all people who are considering being teachers must ask themselves if they are really disposed and prepared to do so. The authors identify four elements that people who wish to immerse themselves in the marvellous task of education must make their own: (i) being a *magister*; which is manifested through self-command of one's conduct and knowledge of the subject one delivers; (ii) *ministry*, as teachers must internally return to what they were to be able to recognise the needs of the people they educate and practise empathy; (iii) to act as *pedagogues*, that is to say, as custodians of those who are entrusted to them; (iv) to act as *maieutic* figures, questioning students so that they come to knowledge of truth. These four elements will, undoubtedly, contribute to the preparation of student teachers, although it is necessary to keep them in mind throughout one's career as a teacher.

The study of the relationship between competence and virtues that authors provide in chapter seven is of great interest. After several definitions of *competence*, they distinguish between technical competences (knowledge and skills from the academic, scientific and technological fields) and personal ones (cognitive, interpersonal and intrapersonal capacities). To achieve an integral education, the person must acquire techniques and, when exercising them, carry out ethical action. Rather than dividing training into personal competences, there should be a tendency towards an integrating concept; what the authors call *personalisation of the formative process*. In this sense, they establish the following relationship between competences and virtues:

(i) Justice as the inclination of the will to give to God and to others what is due to him, so that truth and goodness combine with the interpersonal concept of the student. All of the competences with a social dimension, a dimension of reflection and evaluation of decision making, must be inspired by justice.

(ii) Strength or a firm and constant willingness in the search for good. This virtue must foster the competences of self-criticism and self-evaluation, constance and patience to achieve a strong determination to seek good.

(iii) Temperance, which moderates and ensures balance in the use and attraction of material goods. This is related to the competences of time management and self-knowledge.

There is, however, one supreme, master virtue: prudence. As explained above, this guides and directs our emotions, our intentions, and how we work towards what

is good, the end we wish to achieve. In this way, it pushes or drives away behaviours so that they are useful and directed towards the goal. As a master virtue, it helps decide when to apply one virtue or another, or which is most relevant in a given moment. It discerns good and the means for achieving it. As a corollary to this chapter, the authors note the need for the teacher to seek and exercise these four virtues.

Chapter eight starts with a great truth: "We teach what we are". Education entails a relationship. The educator's way of being has more influence than we might believe, hence the transcendental importance of what is noted in the previous chapters. In this relationship between the educator and the learner, freedom has a special function. Educating can be understood as teaching to be free, conceived as the student's capacity to choose and know how to choose. To do so, it is necessary to combine intelligence and will, functionality and affectivity. As a result, a key distinction emerges between authority and authoritarianism, where the former has degenerated, in part because it has turned into the latter, while the latter is a caricature of the former. Authority emerges from the relationship, from feelings, from the desire to belong, but not through the existence of rules. This could be said to be the difference between *auctoritas* and *potestas*. Teachers, through their virtuous lives, their presence in the classroom, must inspire in their students the love of good and of its implementation.

In the last chapter, number nine, the authors present educational practices that have been carried out in institutions that have opted to focus on an integral education of the student. These practices are of great utility.

We believe this book is essential reading for everyone who has the great fortune of having a vocation (understood as the authors magnificently explain it) for teaching. The way it spells out the concepts and their practical utility is very revealing. Teachers cannot just be transmitters of content; if they were, what would differentiate them from, for example, a *youtuber*? This same thesis is set out masterfully in the novel *La ventana [The window]* by Carmen Guaita, where she skilfully depicts the great influence that teachers have on young people. They can never be replaced by a machine, because with their mere presence, they influence the students. This book does a wonderful job of inspiring this influence to be for the good of the students and for the teachers to become better as people and so as teachers.

Gonzalo Moreno-Ochoa ◼

Ahedo, J. & Arteaga-Martínez, B. (Eds.) (2023).
El tetraedro de la amistad. Estudiante, familia, escuela y sociedad [The tetrahedron of friendship. Student, family, school and society].
Círculo Rojo. 130 pp.

The work *El tetraedro de la amistad. Estudiante, familia, escuela y sociedad* starts with a powerful metaphor that presents friendship as a tetrahedral structure comprising four equilateral triangles in perfect

harmony: society, school, family, and individual. From this perspective, it sets itself the challenging aim of exploring the concept of *friendship* from the educational sphere. This task is tackled through the professional experiences of the different authors, who, chapter by chapter, offer the reader a series of guidelines for understanding how to lead young children in discovering and cultivating the value of friendship.

In the introduction, Pablo Rodríguez, Professor of Molecular Biochemistry and Biology at the Universidad Politécnica de Madrid, offers a scientific analysis of the nature of friendship. He explores how this bond displays itself among primates, where it has been shown to be an essential component for the survival of certain groups. He also examines the evolution of the human species and how this shaped the social structures that permit the formation of bonds of friendship. He concludes that having a solid and balanced social network was crucial for survival in the past and remains a fundamental pillar for social and psychological well-being at present.

In chapter one, Pedro Concejero, a teacher and foster parent, considers two outstanding works of Spanish literature: *Don Quixote* and *El Abencerraje y la hermosa Jarifa [The Abencerraje and the Lovely Jarifa]*, which consider the topic of friendship. As he progresses, he gradually intertwines his personal experience in the field of fostering with the concept of *friendship* and draws meaningful lessons from his own life. In this way, he offers an enriching reflection on the importance of friendship, without neglecting the challenges and hardships that must be overcome if this relationship is to flourish.

The second chapter offers a detailed analysis of the existing academic literature on the construction of friendship from the perspective of attention to diversity. The authors, Jesús Conde-Jiménez and Ángela Martín-Gutiérrez, set out a wide variety of strategies for boosting personal skills that enable the construction of friendship from diversity.

Tania García-Bermejo and Isabel da Mota have written a third chapter that combines the experience of the school environment with university knowledge. To do so, they explore the socio-affective relations that emerge in *casa de niños* preschools. These settings, designed in accordance with the Montessori methodology, can be regarded as microsocieties which, while focussing on individual development, provide a favourable environment for interpersonal development. Through a series of specific experiences, they provide practical illustrations of the content of their study.

The fourth chapter examines how, in a setting where interpersonal dynamics are undergoing a paradigm shift, literature for children and young people can play a crucial role in fostering meaningful relations. So, the authors Sandra Sánchez and María Carmen Torrecilla propose using it as a tool for socio-educational intervention in class, with the aim of promoting competences and critical and reflexive thinking in students while at the same time stimulating dialogue and strengthening their concept of friendship.

In chapter five, David Blanco, a scientific writer and communicator, provides a journey through the most emblematic friendships in the field of science. He explores the relationships between notable scientists, such as Darwin and Wallace or Einstein and Grossman, among others, to show how these friendships not only entailed a personal interchange between these figures, but also how they acted as facilitators of their professional achievements. In this way, he introduces the possibility of integrating, through the study of these friendships, a social dimension that enriches science and makes it more appealing for our students.

Chapter six analyses the method of literacy through photography, which uses photography as a catalyst to consider social relations in more depth. Alfonso Da Silva and Daniela Reyes describe in detail how it can be applied and to what extent it can help students to experience friendship. They also offer a detailed list of guidelines that can help us in this process.

In chapter seven, Silvia Morcillo, a counsellor and educational psychologist, proposes a detailed analysis of the nature and manifestation of friendship in the early stages of life. Instead of adopting simplistic positions that only centre on an egocentric conception of this stage, Morcillo presents multiple pieces of evidence to prove the existence of friendship in these ages. She underlines the crucial importance of these moments in constructing future friendships, as it is during early childhood that we start to internalise values and form our identity,

fundamental aspects for building bonds in the future. To conclude, the chapter proposes a series of guidelines designed to support the development of friendship in these early stages.

María José Bautista-Cerro and María José Díaz explore the importance for our planet of cooperation in chapter eight. In a context of socio-ecological crisis, they underline how the values of sustainability and friendship can be fundamental for creating alternative spaces that foster a more sustainable society in both social and ecological terms. In this sense, they offer guidelines that can be implemented in educational centres and which can, ultimately, contribute to strengthening the social fabric.

In chapter nine, Paola Perochena analyses the crucial role of school counselling in values education. From a perspective of educational psychology, she explores the development of values and friendship and translates the theoretical postulates from the literature into concrete strategies. These strategies guide the reader to promote the development of solid friendships and healthy social relations that help young people to experience true friendship.

The book concludes with a provocative chapter by Josu Ahedo, who poses a fascinating question: what does learning to love a friend mean? Through a detailed analysis of the challenges and opportunities that friendship affords us, Ahedo offers a series of guidelines that facilitate the comprehension of friendship from

three fundamental perspectives: affectivity, the human essence, and personal being.

As a whole, *El tetraedro de la friendship. Estudiante, familia, escuela y sociedad* provides an in-depth, multidisciplinary exploration of the value and nature of friendship from various academic and practical perspectives. The collaboration between professionals and academics provides a wealth of focusses from theory to practical implementation in educational and social settings. Accordingly, the book not only invites us to reflect on the importance of friendship in the life of individuals and in society in general, but it also provides valuable tools to approach it effectively in different contexts. From scientific analysis to practical strategies, each chapter offers ideas and suggestions that enable readers to reconsider how they can cultivate meaningful relations of friendship and foster values such as cooperation, solidarity and sustainability.

This work also stands out for the variety of focusses and the depth with which it approaches the topic, which make it unmissable reading for researchers, educators, parents, and anyone with an interest in understanding and promoting friendship as a fundamental pillar of human development. *El tetraedro de la amistad* not only challenges us to reflect on our own concepts and our perceptions about friendship, but it also provides us with a practical guide to foster more significant and enriching relationships in our lives and communities.

Laura Guerrero-Puerta ■

Martín-Gutiérrez, Á. (2020).
La escuela y su entorno en la sociedad del conocimiento. Un estudio focalizado en la Formación Profesional [The school and its environment in the knowledge society. A study focussed on Professional Training]. Octaedro. 180 pp.

Contemporary education is in a constant process of adaptation and reconfiguration in response to the challenges of a society that is undergoing a rapid transformation. The author of this pedagogical work, *La escuela y su entorno en la sociedad del conocimiento. Un estudio focalizado en la Professional Training*, tackles these challenges from the perspective of Professional Training (PT) in Andalusia, Spain. The book contains an in-depth examination of the crucial relationship between PT educational centres and their surroundings, and underlines the importance of social capital and interdisciplinary collaboration for meeting current work demands and promoting an education that is attuned to the needs of the knowledge society.

Through a detailed analysis, Martín-Gutiérrez examines the dynamics of collaboration that characterise PT centres and identifies their strengths and areas for improvement. In this context, the work presents a consistent theoretical framework, backed by up-to-date and relevant studies. It also proposes a methodological design for evaluating the current collaboration practices in these centres. Its pages take the reader on an in-depth exploration of PT, from its historical and regulatory roots to its current challenges

and proposals for improvement, all of this enriched with empirical data and critical reflections. This review offers an overview of the principal topics and findings presented in the eight chapters of the book. Its aim is to provide a general overview of their content and their relevance in the field of education.

The first chapter develops the approach and rationale of the research. It explains how new demands in the educational sphere emerge from the rapid changes that society currently faces. In this context, participation and cooperation are key elements for improving educational institutions. The author focuses on analysing the relationships of collaboration that PT centres establish with their surroundings, as these ties are fundamental for meeting current employment demands. So, the principal objective of the study is to identify the relations of collaboration that Andalusian PT centre maintain with their surroundings from the perspective of the members of the management team.

The second chapter presents the conceptual framework of the theory of social capital and its link to the knowledge society. Social capital refers to the networks, norms and trust that facilitate coordination and cooperation for mutual benefit. Similarly, the origins of the concept through authors such as Hanifan, Jacobs, Bourdieu, and Putnam are described. Different types of social capital are also listed, such as bonding, bridging and institutional. Finally, there is a reflection on the role of education and PT in generating social capital, at an individual and a communi-

ty level. Through theoretical contributions and analysis, it is concluded that it is necessary to create networks between educational institutions, communities and businesses to promote the development of educational centres.

The third chapter of the work tackles the concept of *educational collaboration* and its importance for improving institutions. It distinguishes between internal and external or inter-institutional collaboration. Within this latter category, different fields are described, such as families, businesses, unions, and employment or counselling services or associations, among others, and the role of management teams in establishing these relations is emphasised. Similarly, concepts such as *learning communities*, *networks* and *centres' openness to their surroundings* are analysed. From all of them, it is deduced that collaboration networks are key to fostering a harmonious environment in the educational centre. Also that teachers can perform effectively in their instructional management and so offer a quality educational service.

Chapter four analyses the development of PT in the European, Spanish and Andalusian context. It reviews regulatory benchmarks such as the LOGSE (Organic Act regarding the General Organisation of the Educational System), LOCFP (Organic Act regarding Qualifications and Professional Training), LOE (Organic Education Act), and LOMCE (Organic Act Regarding the Improvement of the Quality of Education). It also describes initiatives for European participation in PT. Moreover, this chapter explains that PT seeks to link

itself more to the productive environment and the job market. This requires closer collaboration between the centres and other agents such as businesses, unions, employment services, etc. Finally, the laws of autonomous regions such as the LEA (Andalusian Employment Act) and the Andalusian PT act are also analysed, which underline the need to adapt the offer to the local demands and to tighten links between the centres and their settings.

Chapter five presents the methodological design of the study. The phases of reflection, planning, execution, analysis and dissemination are described. The research population comprised 383 PT education centres in Andalusia, from which a stratified random sample of 193 centres was obtained.

A closed questionnaire was administered to the management teams. Its content was validated through expert judgement, and its construct through factor analysis. The questionnaire covered the relevance, agents, characteristics, consequences, and prospects of the relationships of collaboration. The reliability, calculated using Cronbach's alpha, was 0.964.

The results of the questionnaire form part of chapter six. To calculate them, a descriptive analysis of frequencies and percentages was performed. From this, it should be noted that 98% of the management teams placed importance on relations of collaboration, which are principally established with workplace training companies, external companies, local institutions, and other educational centres. In

these spaces for collaboration, objectives relating to students and to links with the productive environment are prioritised. Coordination by the management team facilitates these relations. As consequences, they identified the improvement in students' competences and the building of bridges between professional training and the reality of professional practice. Almost all of them consider it necessary to dynamise the relationships of collaboration.

In chapter seven, Andalusia's professional training centres are classified into four models depending on how developed their collaborative relationships with their surroundings are. Through quantitative techniques, clusters are identified, from incipient collaborations (model A) to consolidated networks (model D). While it is positive that all of the models value this collaboration as crucial, only 2.11% achieve the most advanced stage. This suggests that building solid ties with the surroundings continues to be a challenge. It is important to analyse what organisational or cultural barriers hinder development between models. Another contribution is the proposal of strategies to strengthen collaboration, both internally (training, joint projects) and externally (agreements, networks). Research into its real impact is required to facilitate the transition between stages.

In summary, the diagnosis of the collaborative models is an important contribution. However, it is necessary to consider in depth the obstacles that most centres face in consolidating these relations and determining effective strategies to promote their continued development. Only in this way

can professional training respond flexibly to the demands of society and the productive sector (Rodríguez *et al.*, 2022).

Chapter eight sets out the study's conclusions. These underline the importance that management teams place on collaborative relationships, their principal characteristics, the need to dynamize them, and the models identified. Theoretical, methodological, and practical implications are then proposed. Guidelines are also offered for strengthening relations between the government, management teams, and actors from the surrounding area. Finally, future lines of research are suggested, such as studying other perspectives, analysing PT-business relations, and the role of networks and ICT.

In conclusion, each of the chapters relates to the approach and theoretical framework of the research. Throughout this book, the importance of collaborative relations between PT centres and their surroundings are underlined and the methodological aspects, results, discussion, and conclusions of the research are considered. Furthermore, there is an overview of collaborative relations in Andalusian PT centres, their characteristics, and the prevailing models.

This review synthesises the principal findings and makes it possible to develop an understanding of the content and contributions of the study.

Ángel-Freddy Rodríguez-Torres ∎

References

Rodríguez, Á. F., Medina, M. A., Tapia, D. A., y Rodríguez, J. C. (2022). Formación docente en el proceso de cambio e innovación en la educación. *Revista Venezolana de Gerencia, 27* (8),1420-1434. https://doi.org/10.52080/rvgluz.27.8.43

Table of contents of the year 2024

Articles

709 EV

rep

711 EV

Book reviews

Ahedo, J. y Arteaga-Martínez, B. (Coords.) (2023).
El tetraedro de la amistad. Estudiante, familia, escuela y sociedad [The tetrahedron of friendship. Student, family, school and society]
(Laura Guerrero-Puerta).

n. 289, pp. 700-703

Arthur, J. y Fernández, V. (2023).
En vista del fin excelente: la educación del carácter en las universidades [In view of the excellent goal: Character education in universities]
(Gonzalo Moreno-Ochoa).

n. 289, pp. 697-700

Berkowitz, M. W. (2022).
Modelo PRIMED para la educación del carácter: seis principios esenciales para la mejora escolar [PRIMED for character education: Six design principles for school improvement]
(Unai Buil Zamorano).

n. 287, pp. 180-182

Bernal-Guerrero, A. (Coord.) (2022).
Identidad emprendedora. Hacia un modelo educativo [Entrepreneurial identity: Towards an educational model]
(Jesús Conde).

n. 288, pp. 449-452

Colom, T., Sarramona, J., y Vázquez, G. (2023).
Reflexión y práctica pedagógica [Pedagogical reflection and practice]
(Isabel Alvarez-Canovas y Joana Ferrer-Miquel).

n. 287, pp. 179-180

Fuentes, J. L., Fernández-Salinero, C., y Ahedo, J. (Coords.) (2022).
Democracia y tradición en la teoría y práctica educativa del siglo XXI [Democracy and tradition in educational theory and practice in the 21st century]
(Marta Ambite Pérez).

n. 288, pp. 456-458

Martín-Gutiérrez, Á. (2020).
La escuela y su entorno en la sociedad del conocimiento. Un estudio focalizado en la Formación Profesional. [The school and its environment in the knowledge society. A study focussed on Professional Training]
(Ángel-Freddy Rodríguez-Torres).

n. 289, pp. 703-706

rep

713 EV

Rivas, S., y García-Diego, H. (Eds.) (2022).
Escenarios de aprendizaje. Diálogos entre arquitectura, diseño y educación [Educational settings: Dialogues between architecture, design, and education]
(Carmen María Basanta Vázquez).

Urpí, C. (Ed.) (2022).
Creatividad y bienestar [Creativity and well-being]
(Alicia Encío).

Instructions for authors

A. Purpose of the journal

Revista Española de Pedagogía was created in 1943 and its search for excellence has always distinguished itself. It has been the first journal of pedagogical research in Spanish that has been indexed in the most relevant international databases. It accepts only original, high quality submissions from anywhere in the world that help advance pedagogical knowledge, avoid mere opinion polls, and are of general interest. Articles must follow commonly accepted ethical criteria; in particular, in cases of plagiarism and falsification of data, the author will be penalized by the rejection of their submissions. Articles with more than three authors will only be accepted if a reasoned explanation is provided, and in any case, the intellectual collaboration of all the signatories must be certified, not just data collection. Three issues a year are published.

B. Languages used in the journal

REP publishes all scientific articles and bibliographic reviews in Spanish and English.

When an article is accepted for publication and in order to guarantee the use of correct academic language in both languages, an agreement will be reached with the authors for the translation of their article into English or Spanish. If necessary, the translation will be made by professional experts who are native speakers of each language according to the conditions described in **H. Article Processing Charges (APCs)**. All contents of the original article, including tables and graphs, must be translated.

Texts cited in the article that were originally published in Spanish, even if they were later published in an English translation, must also be included in their original language. In this way, translators will not have to translate these texts again. In particular, it is preferable for a classic text to be cited with both versions: that of its original and that of the printed translation.

C. Requirements of originals

C.1. The publication of research articles must be in accordance with the *Publication Manual of the American Psychological Association* 7th Edition, 2020, (www.apastyle.org). Here are some basic points which must be strictly followed by the authors.

1) The length of the contributions, including all sections, will be between 6000 and 7500 words, using the Times New Roman font.

2) Articles should be submitted following the structure and formats indicated in the template that can be found on the journal's website (https://www.revistadepedagogia.org/rep/plantilla_articulo_eng.docx).

3) In cases where authors have compound names or use more than one last name, such as Hispanic authors, they should be connected with a hyphen. Example: María-Teresa Calle-Molina.

4) The authors must indicate the role of each one using **CREDIT taxonomy** (example available in the template).

5) 6 to 8 keywords should be included.

6) Following the APA model, the References list will be at the end of the article, in alphabetical order by surname, naming all the authors up to a maximum of twenty, with the second line indented.

The translation into English or Spanish should be included in square brackets next to the original title of the publications, since in the Spanish version of the article, the Spanish translations of the titles of the works published in English will be provided. **DOI of publications should be always included whenever possible.**

Some examples are given below:

• **Books:**

Genise, N., Crocamo, L., & Genise, G. (2019). *Manual de psicoterapia y psicopatología de niños y adolescentes [Manual of Psychotherapy and Psychopathology of Children and Adolescents].* Editorial Akadia.

• **Journal articles:**

Siegel, H. (2002). Philosophy of education and the Deweyan legacy. *Educational Theory, 52* (3), 273-280. https://doi.org/10.1111/j.1741-5446.2002.00273.x

• **Chapters in multiauthor books:**

Mendley, D. M. (2005). The research context and the goals of teacher education. In M. Mohan & R. E. Hull (Eds.)*, Teaching Effectiveness* (pp. 42-76). Educational Technology Publications.

• **References to web page:**

Guarino, B. (2019, January 3). How will humanity react to alien life? Psychologists have some predictions. *The Washington Post*. https://www.washingtonpost.com/news/speaking-of-science/wp/2017/12/04/how-will-humanity-react-to-alien-lifepsychologists-have-some-predictions

U.S. Census Bureau. (n.d.). *U.S. and world population dock. U.S.* Department of Commerce.

Retrieved July 3, 2019, from https://www.census.gov/popclock/

7) References in the body of the article are written in an abbreviated way that differs from what is used in the Reference list. Specifically, if the reference is a direct quotation, the text must be enclosed in quotation marks and, usually at the end, the author's last name, year and page number are placed in parentheses: "(Taylor, 1994, p. 93)". If it is not a direct quotation, and so is not enclosed in quotation marks, the page number will be omitted: (Taylor, 1994). When the author's name is given in the text he/she will not be included in the parenthesis: "According to Taylor (1994, p. 93), culture ..." When an idea is supported by several authors, they will be separated by semicolons: "(Taylor, 1994; Nussbaum, 2012)".

To quote several works by one author, only the years will be added after the author, with letters added if it is necessary to distinguish between publications from the same year: "(Taylor, 1994, 1996a, 1996b)".

When citing works by 3 or more authors, only the first one is cited followed by "et al".

Textual quotes will be written in-line if they have fewer than 40 words. If the quotation has 40 words or more, it will be placed in a separate paragraph, without quotation marks, indented by 0.5 cm and in the body text style in a typeface one point smaller. Following the quotation, the author, the year and the page are added in parentheses. The material quoted is reproduced textually, including spelling and punctuation.

Other authors' texts will be quoted following the criterion of consulting the originals that are written in those languages and using their official translation when such text has also been edited in the other language. If this official translation is not available, the quoted text will be offered to the readers translated by the author of the article (noting that the translation belongs to the author of the article), or by the sworn translator hired by the journal.

The use of endnotes will be limited. They must have correlative numbering, using the automatic system in Word and they will be placed after the body of the article and before the References that list everything cited in the text.

8) To highlight a word, italics will be used. Underlining or bold should not be used.

9) The number of lists, diagrams, tables and figures in the text should be limited. These will be called Tables or Figures. In any case, they must be where they should be in the article. In tables, columns should be aligned using tabs (only one tab per column). When quoted in the text (e.g., "as we see in Figure 1 on core subjects"), only the first letter will be capitalized, while at the top of the Table or Figure the whole word will be in small caps, in 12 point capital with Arabic numerals, followed by a point, writing the title in normal text.

The text within the table will be written in the same typeface as the normal text and in 9 point. The source of the table or figure will be placed below it, without a space of separation, stating the Source, colon, surnames, comma and year.

Graphs and tables, in addition to appearing where they should in the article, have to be sent in their original editable format whenever possible. Images should always be sent in high resolution (300 dpi).

10) Equations will be centered, separated from the main text by two lines. They should be referenced in the text, stating the number of the equation; therefore, they will be accompanied by Arabic numerals, aligned to the right and in parentheses in the same line.

11) The article will conclude with a list of the bibliographical references of all the works cited, except for the works cited whose authors include one of the authors of the article. In these cases, these works will be listed in the version with names of the authors, while in the anonymous ones they will not be included in the references, although they will appear in the text, where they will appear as follows: "(Author, 2022, p.

39)". Citation of publications belonging to journals or publishers considered "predatory", i.e., those that lack a rigorous and quality scientific evaluation system (e.g., double-blind peer review) and whose main purpose is not to disseminate knowledge but to obtain an economic profit by charging publication fees to authors. Lists of predatory publishers and journals can be consulted at: https://beallslist.net/

12) Finally, a brief biography of the authors should be included, of a maximum of ten to fifteen lines, which should mention their ORCID and the main aspects of their academic career, current academic situation and the university where they obtained their higher academic degree.

13) Authors of published works that have been carried out with research data including the sex variable are encouraged to report whether the conclusions have taken into account possible differences between sexes.

C.2. In addition to research articles, the **Revista Española de Pedagogía** wishes to keep up to date by publishing, in various formats, other works and relevant information in pedagogical science. For this reason, it publishes reviews of books, current news, brief commentaries on educational problems, readers' comments on articles published in the last year, etc. The reviews, always on recent books from relevant publishers, will be between 1200 and 1700 words. They will be headed by the book's details as follows:

Villardón-Gallego, L. (Coord.) (2015). *Competencias genéricas en educación superior [Generic competences in higher education]*. Narcea. 190 pp.

Commentaries will be of moderate length. The analysis of published articles will be sent, from the journal, to the author of the analysed article, so that he/she can prepare a response.

D. Policy on the use of artificial intelligence (AI) in articles

Authors must follow the AI use policy established by Revista Española de Pedagogía and declare

compliance with the following sections before submitting their articles.

1) Authorship of the article:

- Author(s) cannot cite AI as the author or co-author of the submitted articles.

2) Use of AI in the writing process:

- The authors may only use generative AI or AI-assisted technologies to improve the language and readability of this article.

- If AI is used, authors must cite it in the References section according to the use of APA standards in force in the journal.

3) Use of Language Multimodal Model (LMM) or Large Language Model (LLM) in the development of the article:

- The authors are responsible for reviewing and validating the AI-generated information.

- The authors should indicate and document the use of LLM or LMM in the Methods section.

4) Use of AI-generated images and videos in the article:

- The use of AI-generated images and videos in articles is not permitted.

5) IA Policy Compliance:

- In the case of non-compliance with the IA policy, the journal may reject (pre-publication), retract (post-publication), or publish an editorial notice on the article.

The acceptance of **this declaration is mandatory** if the authors wish to publish it in the journal.

E. Submissions

Submission of papers is open on a permanent basis. Special deadlines will be established for publicly announced monographic issues.

All papers must be submitted through the journal's web platform by clicking on the following link: Enviar artículo | Submit Article

After creating the user, the system will guide you through a series of pages in which you will be asked to accept the Submission agreement, the journal's policies, the requirements for papers submitted to the journal, and to enter the data related to the authors and the submitted article. Only papers that comply with these policies and journal requirements will be considered for evaluation.

The system accepts Word or RTF files. The system will automatically generate a PDF for you. Images or graphics should be placed in the article in their proper place in the highest possible quality. If the images are not of sufficient quality (300 dpi), they will be requested again from the authors once their article has been accepted for publication. Tables should be placed in the article in the appropriate place and be editable.

The corresponding author will receive an automatic notification confirming receipt of the article. This notification will indicate the link through which you will be able to access your article on the journal's web platform and make any modifications or send new files that may be necessary during all the evaluation and editing process of the article.

F. Submissions evaluation and editorial processing times

The editors will make an initial assessment of the article based on the suitability of the proposals to the editorial criteria of the **REP**. Authors can consult more information on the type of articles accepted by the journal at the following link: Scope. The result of this assessment will be notified within a maximum period of one month. If a positive first evaluation is received, the paper will be peer-reviewed. Authors will receive the result of the second evaluation within approximately three months.

Each article submitted will be subject to the journal's editorial decision process. The journal is under no obligation to publish the article.

The deadline established for the completion of the evaluation process is four months, counting from the notification of receipt of the article. At the end of this period, the author will normally be informed of the final result of the evaluation. Authors can consult more information on the evaluation procedures followed by the journal at the following link: Submissions evaluation and editorial processing times. An author whose article has not been selected may resubmit other papers at a later date.

Accepted articles will begin the editing process (translation, style correction, layout, etc.), to be subsequently included in the corresponding issue, according to the decision of the editorial direction. Once the editorial process has been completed, the preliminary layout of the text will be sent to the authors for final revision and approval. The editing process usually takes a maximum of two months.

The average time between the receipt of an article and its publication is six months.

Publication of articles does not entitle the author to any remuneration.

G. Publication costs

REP provides diamond open access. Publication is free and open with no costs to authors or readers.

H. Article Processing Charges (APCs)

In order to guarantee the quality of published scientific articles, we follow a policy of only accepting professional translations from accredited translators or translations from authors whose mother tongue is Spanish or English or who have a high level of proficiency in these languages.

Once an article has been approved for publication in the **REP**, in cases where a professional translation into English or Spanish is required, authors must pay for the translation under the following conditions:

- If it is an article supported by a research grant, or similar, the invoice will be paid in full by the body supporting the research. The total cost will generally depend on the number of words translated. If this institution has an accredited professional translation service, translations carried out by the funding body will also be accepted.

- In the case of an article without research or publication support, the author will pay the maximum amount of 400 euros (including 21% VAT), and the journal will be responsible for the rest of the quoted translation costs, which will depend on the number of words translated.

The costs of translating bibliographical reviews or other non-scientific content published in the journal in Spanish and English will be fully covered by the journal.

I. Dissemination of published papers

Once the papers have been published in the **Revista Española de Pedagogía**, authors can contribute to dissemination tasks, both by supporting the ones that the journal itself carries out and by their own initiative.

The **Revista Española de Pedagogía** has profiles on the main social networks (Facebook, Twitter and LinkedIn), where it disseminates the papers it publishes, consequently we recommend that authors follow the journal on these networks and share their publications.

https://www.facebook.com/revistadepedagogia

https://twitter.com/REPedagogia

https://www.linkedin.com/company/revista-espanola-de-pedagogia

To assist in the dissemination of their articles, once approved for publication, authors will be asked to send:

- Two key summary sentences of their article of a maximum of 180 characters for use on social network X.

- A summary of your article of about 90 words for use on Facebook and Linkedin.

- A video of approximately 1 minute in length, in horizontal format, summarizing the main ideas developed in the article for use on our social networks would also be appreciated.

- Our journal is also part of the academic blog Aula Magna 2.0 (http://cuedespyd. hypotheses.org/), where entries on topics of interest for educational research are published periodically, as well as reviews of articles, which contribute to its dissemination. Aula Magna 2.0 publishes an entry dedicated to an article of the **REP** for each published issue, for which the authors will be asked to provide a longer summary, of between 600 and 1500 words, in a langua-

ge accessible to the general public and a high-resolution photograph.

Authors are also encouraged to deposit or disseminate accepted articles in:

- Institutional repository of their university and public repositories (SSRN, Zenodo, etc.).

- Google Scholar, ORCID, Dimensions, PlumX, etc.

- Scientific social networks.

- Social networks (Facebook, Twitter, LinkedIn, etc.).

- Personal or institutional website, blog, etc.

It is required that all the bibliographic data of the published work be detailed in these publications.

Each author should consider using the most effective means of disseminating their article, obtaining citations and contributing to the advancement of pedagogical knowledge.