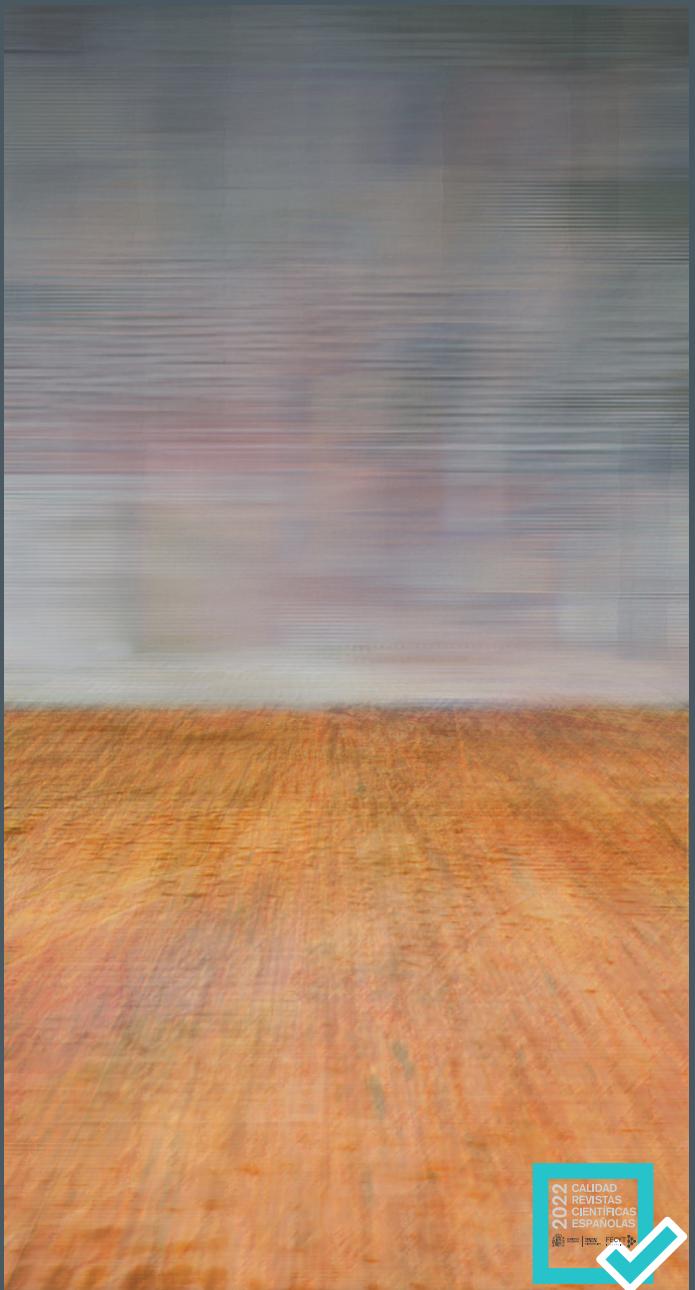


revista de EDUCACIÓN



Nº 401 JULY-SEPTEMBER 2023



GOBIERNO
DE ESPAÑA

MINISTERIO
DE EDUCACIÓN
Y FORMACIÓN PROFESIONAL



revista de
EDUCACIÓN



Nº 401 JULY-SEPTEMBER 2023

revista de EDUCACIÓN

Nº 401 July-September 2023

Quarterly Journal

Starting year: 1952



MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL
SECRETARÍA DE ESTADO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL

Instituto Nacional de Evaluación Educativa
Paseo del Prado, 28, 4.^a planta
28014 Madrid
España

Edita
© SECRETARÍA GENERAL TÉCNICA
Subdirección General de Atención al Ciudadano, Documentación y Publicaciones

Catálogo de publicaciones del Ministerio: sede.educacion.gob.es
Catálogo general de publicaciones oficiales: publicacionesoficiales.boe.es

Edición: 2023
NIPO pdf: 847-19-002-9
NIPO ibd: 847-19-001-3
NIPO html: 847-21-218-4
ISSN línea: 1988-592X
ISSN papel: 0034-8082
Depósito Legal: M.57/1958

Diseño de la portada: Dinarte S.L.
Maqueta: Glaux Publicaciones Académicas S.L.U..

BOARD OF DIRECTORS

CHAIR

José Manuel Bar Cendón
Secretario de Estado de Educación

Clara Sanz López
Secretaria General de Formación Profesional

MEMBERS

Liborio López García
Subsecretario de Educación y Formación Profesional

Mónica Domínguez García
Directora General de Evaluación y Cooperación Territorial

Santiago Antonio Roura Gómez
Secretario General Técnico

Carmen Tovar Sánchez
Directora del Instituto Nacional de Evaluación Educativa

Clementa López Fernández
Subdirectora General de Atención al Ciudadano, Documentación y
Publicaciones

Purificación Llaquet
Subdirectora de Cooperación Territorial e Innovación Educativa

Lucio Calleja Bachiller
Subdirector General de Ordenación Académica

EDITORIAL BOARD

DIRECTOR

Carmen Tovar Sánchez

EDITOR IN CHIEF

José Luis Gaviria Soto

ASSISTANT EDITOR

David Reyero García

MEMBERS

Antonio Cabrales Goitia (University College London); Caterina Casalmiglio (Universitat Autònoma de Barcelona); Antonio Lafuente García (Consejo Superior de Investigaciones Científicas); Leocadio López-Ocón Cabrera, (Consejo Superior de Investigaciones Científicas); Carlos Marcelo García (Universidad de Sevilla); Miquel Martínez Martín (Universitat de Barcelona); Francisco Michavila Pitarch; (Universidad Politécnica de Madrid); Juan Manuel Moreno Olmedilla (Banco Mundial); Clara Eugenia Núñez (Universidad Nacional de Educación a Distancia); Lucrecia Santibáñez (Claremont Graduate University); Denise Vaillant (Universidad ORT, Uruguay); Pablo Zoido (Banco Interamericano de Desarrollo)..

EDITORIAL TEAM

Technical editors: Sabrina Gallego Verdi y Juan Carlos Mejía

Collaborator: Ruth Martín Escanilla

SCIENTIFIC ADVISERS

International

Aaron Benavot (State University of New York, SUNY-Albany); Abdeljalilb Akkari (Universidad de Ginebra); Mark Bray (University of Hong Kong); José Joaquín Brunner (Universidad Diego Portales, Chile); Dirk Hastedt (Executive Director, International Association for the Evaluation of Educational Achievement, IEA); Andreas Schleicher (Director, Directorate for Education and Skills, OCDE).

National

Teresa Aguado Odina (Universidad Nacional de Educación a Distancia); Margarita Bartolomé (Universitat de Barcelona); Antonio Bolívar (Universidad de Granada); Josefina Cambra Giné (Colegio de Doctores y Licenciados); Anna Camps i Mundó (Universitat Autònoma de Barcelona); César Coll Salvador (Universitat de Barcelona); Agustín Dosil Maceira (Universidad LiberQuaré); Gerardo Echeita Sarrionandia (Universidad Autónoma de Madrid); Juan Manuel Escudero Muñoz (Universidad de Murcia); Mariano Fernández Enguita (Universidad Complutense de Madrid); Joaquín Gairín Sallán (Universitat Autònoma de Barcelona); José Luis García Garrido (Universidad Nacional de Educación a Distancia); Daniel Gil Pérez (Universitat. de València); Fuensanta Hernández Pina (Universidad de Murcia); Carmen Labrador Herraiz (Universidad Complutense de Madrid); Miguel López Melero (Universidad de Málaga); Elena Martín Ortega (Universidad Autónoma de Madrid); Rosario Martínez Arias (Universidad Complutense de Madrid); Inés Miret (Neturty S.L., Madrid); Juan Ignacio Pozo (Universidad Autónoma de Madrid); Joaquim Prats Cuevas (Universitat de Barcelona); Manuel de Puelles (Universidad Nacional de Educación a Distancia); Tomás Recio Muñiz (Universidad de Cantabria); Luis Rico Romero (Universidad de Granada); Juana M. Sancho Gil (Universitat de Barcelona); Mercedes Vico Montelola (Universidad de Málaga); Antonio Viñao Frago (Universidad de Murcia); Jesús García Laborda (Universidad de Alcalá)..

Index

Research

JOSÉ SANTIAGO ÁLVAREZ MUÑOZ, M ^a ÁNGELES HERNÁNDEZ PRADOS & JUAN ANTONIO GIL NOGUERA: Family perception of homework as a function of the responsible agent	1
JOSÉ FRANCISCO MARTÍN ALONSO & ISABEL MUÑOZ SAN ROQUE: The paradox of learning to learn development: adapting a scale in secondary school students	25
BELÉN SÁENZ-RICO DE SANTIAGO, M ^a DEL ROSARIO MENDOZA CARRETERO, RAÚL GARCÍA MEDINA & MERCEDES SÁNCHEZ SÁINZ: Challenges in teaching practices for the incorporation of sustainable development approach in Compulsory Secondary Education	53
CSABA SZABÓ, CSILLA ZÁMBÓ, ANNA MUZSNAY, JANKA SZEIBERT & LÁSZLÓ BERNÁTH: Investigating the efficacy of retrieval practice in university mathematics	79
IRENE SÁNCHEZ-SAN-JOSÉ, DANIEL ENGUÍDANOS, MANUEL IGLESIAS-SOILÁN, PATRICIA ARIZA, JAVIER AROZTEGUI & JUAN FERNÁNDEZ: An alternative teaching and learning methodology at university to counter academic boredom	97
ARIADNA HERNAIZ-SÁNCHEZ & ANDREA BÄDER-GILABERT: Influence of Physical Education on physical activity levels in adolescence. A systematic review	123
MÍRIAM ABIÉTAR LÓPEZ, JOAN CARLES BERNAD I GARCIA, ANA ISABEL CÓRDOBA IÑESTA, ELENA GIMÉNEZ URRACO, ESPERANZA MERI CRESPO & ALMUDENA ADELAIDA NAVAS SAURIN: Differential factors in Vocational Education and Training itineraries: A longitudinal study	147
PEDRO JOSÉ CARRILLO-LÓPEZ: Diet quality and academic performance in schoolchildren: the moderating role of weight status	171

FRANCISCO QUIÑONEZ TAPIA & MARÍA DE LOURDES VARGAS GARDUÑO: Mastery-Approach Goals in Secondary Education Students	193
FRANCISCO LÓPEZ RUPÉREZ & ISABEL GARCÍA GARCÍA: A quantitative SWOT analysis for Spanish education.....	219
MANUEL ALFREDO MARCOS BALABARCA: Factors influencing the development of reading competencies in Peruvian students in second-grade of secondary education.....	251
CHUAN-CHUNG HSIEH, IMAM GUNAWAN & HUI-CHEH LI: A Bibliometric Review of Instructional Leadership Research: Science Mapping the Literature from 1974 to 2020.....	281
Reviews	311



Research section

Family perception of homework as a function of the responsible agent

Percepción familiar de las tareas escolares en función del agente responsable

<https://doi.org/10.4438/1988-592X-RE-2023-401-581>

José Santiago Álvarez Muñoz

<https://orcid.org/0000-0002-3617-215X>

Universidad de Murcia

M^a Ángeles Hernández Prados

<https://orcid.org/0000-0002-9740-6175>

Universidad de Murcia

Juan Antonio Gil Noguera

<http://orcid.org/0000-0003-3384-0607>

Universidad de Murcia

Abstract

The controversies caused by the subject of homework, both in the family and in educational centres, place it as one of the focuses of interest for educational research. It focuses on the diversity of support agents to which the student can turn, an aspect that has been scarcely dealt with, and considering that they can influence the homework completion process. Objective: To determine the perception of benefits, emotional exhaustion caused, the child's competences and the organisation of homework according to the family or non-family member who is the support agent. Method: A non-experimental, correlational, quantitative-cross-sectional research study was carried out with the aim of analysing this condition. For this purpose, the questionnaire "Coping with homework in the family" was applied to a total of 1787 families of infant (15.1%), primary (48.7%) and secondary (36.2%) school students. Results: The results revealed that the

person who mainly accompanies, in the completion of homework, is the mother and that, based on the different support figures (fathers, mothers, brothers and sisters or external professional), significant differences are established in relation to the organisation with homework, the perception of the benefits of homework, the competences that students have to cope with homework and the emotional wear and tear that it can produce in the family. Conclusions: The need to educate families so that they see homework as a positive task for the child's learning and as a possibility of bringing the family closer together to strengthen ties instead of causing estrangement is evident.

Keywords: family, homework, parent-child relationship, parent-school relationship, family accompaniment.

Resumen

Las controversias que ocasiona el tema de los deberes escolares, tanto en la familia como en los centros educativos, lo sitúan como uno de los focos de interés para la investigación educativa. Poniendo la mirada sobre la diversidad de agentes de apoyo a los que el estudiante puede recurrir, aspecto escasamente tratado, y considerando que éstos pueden influir en los procesos de realización de las tareas. Objetivo: Determinar la existencia de diferencias significativas en la percepción familiar de los beneficios, el desgaste emocional ocasionado, las competencias del hijo/a y la organización de los deberes escolares en función de la persona (familiar o no familiar) que se encarga del seguimiento y acompañamiento de las tareas escolares de los menores escolarizados. Método: Se llevó a cabo una investigación de metodología no experimental, correlacional de corte cuantitativo-transversal, con el objetivo de analizar esta condición. Para ello se aplicó el cuestionario "El afrontamiento de las tareas escolares en la familia" a un total de 1787 familias de estudiantes de educación infantil (15,1%), primaria (48,7%) y secundaria (36,2%). Resultados: Estos revelaron que la persona que acompaña principalmente, en la realización de las tareas, es la madre y que, en base a las distintas figuras de apoyo (padres, madres, hermanos y hermanas o profesional externo), se establecen diferencias significativas en relación con la organización con las tareas, a la percepción de los beneficios que tienen los deberes escolares, a las competencias que tienen los estudiantes para hacer frente a las tareas y el desgaste emocional que pueden producir en la familia. Conclusiones: Se pone de manifiesto la necesidad de educar a las familias para que estos aprecien las tareas escolares como una tarea positiva para el aprendizaje del menor y una posibilidad de acercamiento familiar en la que estrechar lazos en lugar de ocasionar distanciamiento.

Palabras clave: familia, tareas escolares, relación padres e hijos, relación padres-escuela, acompañamiento familiar.

Introduction

There are different ways of naming it, with "academic tasks" being the most commonly used in the scientific context and "homework" or "school-work" the most common in educational praxis. Conceptually, it is understood as the additional or supplementary homework assigned to students to be developed outside of class that must be guided and controlled by the curriculum (Pérez et al., 2020). Building on other studies, Valiente-Barroso et al. (2020) conceive of homework as "instructional strategies temporarily outside of teacher mediation that students carry out without direct teacher support and attention" (p.152). For Valle et al. (2021), it is not the place, but the time of completion that defines homework, as it is usually done outside school hours.

To talk about homework is necessarily to make an explicit mention of family mediation and intervention. Few school resources are as important in the family context as homework. In fact, homework acts as the main link between the two contexts and is one of the topics most frequently dealt with in family tutorials. Thus, far from encouraging participation, these spaces for communication between family and school are aimed at monitoring study, performance and behaviour in the classroom, especially if this is problematic (García-Sanz et al., 2010; Rodríguez-Ruiz et al., 2019) at the national level, while internationally, a partnership of associationism with one focus on the school and the other on the co-community (Epstein, 1990; Hoover-Dempsey and Sandler, 1995; Mylonakou and Kekes, 2005).

Among the models that analyse homework, as Álvarez-Blanco (2019) points out, the overlapping spheres model by Epstein (1990), the parental involvement model by Hoover-Dempsey et al. (1995), the model of predictors of parental involvement by Grolnick et al. (1997) and the Syn-education model to promote parental involvement in the educational process developed by Mylonakou and Kekes (2005) stand out. From a global perspective, Trautwein et al. (2006) detail the relationships that occur on the learning environment, family behaviour, student characteristics, homework motivation and homework behaviour and their influence on academic performance (Rosário et al., 2018), the latter being the closest to our study.

Family involvement and, more specifically, homework support are two of the essential variables that determine overall school performance

(Martínez-Vicente et al., 2020; Wilder, 2014), and this is perceived as such by the entire school community, including the students themselves (Pérez et al., 2020; Pineda and Fraile, 2020). Focusing on family support for homework and academic performance, there is a great deal of controversy and no clear association has been established. Despite the multitude of studies that recognise that family involvement contributes to school improvement, the recent meta-analysis by Fernández-Alonso et al. (2022) concludes that family help with homework is not a predictor of good academic results, although it is likely that this paradoxical phenomenon is due to the fact that there is greater family involvement when students show poor school performance, learning difficulties, problems in the organisation of study and homework, or low motivation. Similarly, the stage with the highest volume of academic homework in Spain is Primary Education, although it has been noted that homework is of little use in this stage, as well as not being certain whether it has been done independently or whether it has been done by a family member or private teacher (Bailén and Polo, 2016). Even so, Nuñez et al. (2017) highlighted that indirect involvement, through motivation, has a more positive impact on performance than more controlled participation. All of this shows that the subject of homework presents elements that require analysis, study and clarification, despite the large volume of publications on the subject. Who accompanies homework, the type of accompaniment and how the person who accompanies has an impact on the perception of the usefulness or otherwise of homework, are aspects that have yet to be clarified.

Most theoretical models highlight the co-responsibility of parents or guardians in homework, especially in the case of the youngest children, since as they advance in the school system, students should gain in autonomy and assume the main responsibility for their homework (Murillo and Martínez-Garrido, 2013). In contrast, the results obtained by Valiente-Barroso et al. (2020) confirm that students receive more help from their families as they progress through school. However, we find that very few studies have considered, on the one hand, whether or not support is needed, and on the other hand, the role played by the person who assumes the support. In this respect, Hernández-Prados and Gil-Noguera (2022) found that students tend to do their homework alone, autonomously, especially in the case of children whose parents have higher education, while families with a lower academic level prefer the support of external professionals.

The perception of the usefulness of homework is controversial, as some studies highlight the benefits and skills that students should put into practice, attributing an intrinsic positivity to it (Sallee and Rigler, 2008; Tan et al., 2020), while others highlight its difficulties (Kohn, 2013), which is why it has also been questioned by various groups in the school community. Although teachers consider them to be an adequate resource to promote learning, they also point out that parental accompaniment is the most controversial issue, as students must work autonomously (Arias-Redondo et al., 2006), parents perceive the volume of homework to be high, question its effectiveness and usefulness, as it is repetitive, difficult and boring, but still attribute benefits (Feito-Alonso, 2020; Pérez et al., 2020), despite generating chaos in the family climate, conflict, stress and some discomfort (Pressman et al., 2015). All this has sparked an immense social debate, which was highlighted by the awareness-raising campaign launched by the Spanish Confederation of Parents' Associations (CEAPA) on the need to avoid collapsing students' family time with homework (Hernández-Prados and Gil-Noguera, 2018). In fact, it is difficult to balance homework support with extra-curricular activities (Pressman et al., 2015). Studies on peer support in homework have focused mainly on fathers and mothers (Murillo and Martínez-Garrido, 2013; Valiente-Barroso et al., 2020), and to a lesser extent, on the help provided by brothers or sisters or that received by external professionals. In addition, parental accompaniment is conditioned by gender, with a feminisation of this function being observed (Fernández-Freire et al., 2019; Lehner-Mear, 2021; Páez and Zúñiga, 2021; Valdés-Cuervo et al., 2022) and identifying a significant relationship between maternal characteristics and motivation towards tasks, so that an accompaniment focused on autonomous learning and the active role of the child decreases maternal involvement (Valdés-Cuervo et al., 2022). On the other hand, for mothers, school matters are part of the routine tasks of family life, but they are understood as a collective activity in which siblings can also collaborate by learning together to model study habits and time management (Bempechat, 2019). With regard to accompaniment between siblings, Van der Kaap-Deeder et al. (2017) showed that not only parents and mothers are role models for homework, but they also found that brothers and sisters play an important role in the development of autonomy, psychological support, motivation and improved academic performance. Moreover, in the case of

students from immigrant families, compared to their native-born peers, they turn more to their siblings than to their parents for homework help (Bempechat, 2019). On the other hand, older siblings who take on the responsibility for school care and support feel fatigued and stressed, find it more difficult to concentrate and their school performance decreases (Reimer and Hill, 2022).

In view of the above, there is no doubt that we are dealing with a relevant and complex relationship, as there is a multitude and diversity of variables to consider, beyond the quantity or difficulty involved, as well as different sources of information (mainly from students, teachers and families). Although the topic of homework has generated a great deal of research in recent years, a review of previous studies confirms that the variable, who is involved in homework, has been treated in a descriptive rather than a leading role as an element of possible differentiation. Thus, the main purpose of this study was to find out whether there are significant differences in the family's perception of homework depending on the person who is responsible for accompanying the child. This is articulated in a series of specific objectives that demarcate the organisation and presentation of the results:

- To check if there are differences in the difficulty, quantity and modality of homework according to the person who carries out the accompaniment.
- To find out whether the benefits and competences attributed to homework vary significantly according to the support agent.
- To find out whether the limitations and emotional exhaustion of the family depend significantly on the person responsible for the homework accompaniment.

Method

The study corresponds to a quantitative perspective of a descriptive-explanatory, comparative and non-experimental type. The survey technique was applied using a transactional design, i.e., an analysis was carried out at a specific moment in time (Sáez-López, 2017).

Sample

In the present research, 1787 families with children enrolled in a Spanish educational centre in the stages of Infant Education, Primary Education and Secondary Education participated, with 67.55% of the questionnaire being completed by the mothers and 32.45% by the fathers. Both mothers ($M= 42.73$ years; $S.D.= 12.49$) and fathers ($M= 44.70$ years; $S.D.= 6.49$) are slightly over 40 years of age on average. The sample includes families from all the Autonomous Communities, but the Region of Murcia (21.3%), Andalusia (17.1%) and the Community of Madrid (16.6%) are the most representative. More than half of the fathers (50.1%) and mothers (65.5%) have a university education, 10.1% of both genders are foreign nationals and there is a disparity with regard to employment, as 26.6% of mothers, compared to 13.0% of fathers, are unemployed. Finally, with regard to academic variables, on the one hand, 80.1% of the children attend public schools while 19.9% attend private-controlled schools and, on the other hand, almost half of the parents have children in Primary Education (48.7%), with a considerable sample of children also attending Secondary Education (36.2%), while those attending Infant Education (15.1%) are much lower. The only criteria for inclusion in the research were to be a resident in Spain and to have children enrolled in pre-school, primary or secondary education, thus selecting the participating population on the basis of non-probabilistic random sampling.

Instrument

Family accompaniment in homework was measured using an ad hoc questionnaire, designed by means of a systematic process that included a bibliographic search of national and international studies from the last five years indexed in the Dialnet, WOS and Scopus databases and a process of expert validation in which the following phases were followed: definition of the objectives and the evaluation instrument, delimitation of the characteristics for the constitution of the group of experts, carrying out rounds of consultation and analysis of the data obtained. In this way, seven former experts took part, four female teachers in non-university educational stages with a range of experience of five to ten years and three female university lecturers who are experts in educational research

methodology with a working experience of 10 to 20 years. These experts gave a quantitative (Likert-type scale from one to five) and qualitative assessment of each of the parts and all of the items of the initial instrument designed. The comments and assessments made contributed to the elimination of items, modification of feedback, inclusion of inclusive language, inclusion of new items, restructuring of the dimensions and the introduction of new clarifications in the introduction of the questionnaire. After this process, the final research questionnaire entitled: "Coping with homework in the family" was elaborated.

This instrument is made up of 24 items, which were evaluated on a Likert-type scale from one to four according to the following categorisation: one (not at all), two (a little), three (quite a lot) and four (a lot). All the items that make up the questionnaire are grouped into four dimensions:

- Organisation (items 1 to 5): difficulty, quantity, group tasks, individual tasks and digital tasks.
- Benefits (items 6 to 11): evaluates the improvements obtained by the children as a result of homework: learning and performance, interpersonal relationships, study habits, attitude towards school, curiosity and research, and family relationships.
- Competences (items 12 to 19): it asks about the children's capacities for this school performance, these are: knowledge mastery, understanding instructions, technological tools, time organisation, taking responsibility, effort, motivation and emotional management.
- Emotional attrition (items 20 to 24): the degree of discomfort caused to parents is measured from the following descriptors: discomfort, exhaustion, academic inability to help them, family conflict or lack of time.

The validity of the construct was evident from the application of an Exploratory Factor Analysis (EFA) in which Barlett's test of sphericity obtained a perfect value ($p=0.000$) and a KMO coefficient categorised as very high (.910), representing four factors that account for 60.33% of the total variance, obtaining a value of over .500 in all factor loadings. Furthermore, from the Confirmatory Factor Analysis (CFA), according to Kline (2005), an almost perfect goodness of fit is shown ($SB\chi^2=19855.006$; $gl=276$;

$p=.000$; CFI=0.96; TLI=.91; RMSEA=.07; RMR=.06). Finally, regarding the reliability of the instrument, from the use of Cronbach's Alpha parameter, high or very high values are obtained at the global level ($\alpha=.880$) and for its sub-dimensions: organisation ($\alpha=.743$), benefits ($\alpha=.847$), competencies ($\alpha=.911$) and emotional exhaustion ($\alpha=.830$) (Olson and Gorall, 2006).

Procedure

For the composition of the participant sample, communication was established with the educational centres and parents' associations by telematic means through e-mail, establishing contact with 50 in each autonomous community. The message contained a full description of the research (objectives, rights of participants and use of data), the guarantee of anonymous data processing and the complete questionnaire together with the link to invite families to fill it in. A total of 67 schools accepted the invitation by signing an informed consent form assuring the confidentiality of the data and the anonymity of the participant, following the principles set out in point 8.2 of the APA regulations.

Data analysis

Initially, the data were collected through Microsoft Excel and then transferred to the statistical package SPSS version 25.0. Before carrying out the comparison of the variables, all the dependent variables referred to in the four dimensions (24 items) were recoded from a 1 to 4 scale with four categories to a dichotomous variable: on the one hand, those who answered the value one and two (not at all and a little) and, on the other hand, those who selected the value three and four (quite a lot and a lot). This made it possible to obtain the frequencies of all the dependent variables from a new grouping according to the perception of the agent who helps the most in school tasks.

Subsequently, frequencies (percentages) were obtained for the question of who helps with the children's homework. Subsequently, through the normality test, it was determined that the sample does not follow the assumptions of normality (Kolmogorov-Smirnov) and homoscedasticity (Levene's test), so non-parametric statistics were selected. Thus, in order

to compare two nominal categorical variables, the person who helps with homework and the items of the dimensions of organisation, benefits, competencies and emotional exhaustion, the Chi-Square test was used with a level of significance $\alpha= 0.05$. Finally, in order to determine the intentionality of the associations found, Cramer's V parameter was applied.

Results

Table I shows the statistical data regarding the frequencies of the categories of the main study variable for the identification of significant differences: Who helps with homework? The participating families point to mothers as the main agent of support (64.9%), with a considerable distance from the rest. Those families who consider that their children do not need help (17.3%) are higher than those who indicate the father (12.5%). The proportion of those who count on the collaboration of siblings (2.5%) or an external professional (2.7%) is minimal compared to the rest of the percentage values previously mentioned.

TABLE I. Frequency of agents providing homework help

Who helps?	Frequency	Percentage (%)
Father	224	12.5
Mother	1160	64.9
Siblings	44	2.5
External professional	49	2.7
No help required	310	17.3

Source: Compiled by author

The results of the relationship between the helping agent and the consideration of the families with respect to the items of the homework organisation dimension, shown in Table II, show significant differences with respect to the frequency of group homework ($\chi^2 = 53.993$, $p=.000$) and digital homework ($\chi^2 = 34.628$, $p=.001$), being higher than the minimum level of significance .050 in the rest of the items. In the first of the cases, with respect to group homework, depending on the consideration of the support agent, a low consideration of the amount of this type of homework is identified in all the agents, although this is less pronounced

in those families who identify the sibling as the main support. In relation to digital homework, those families who perceive the parent as the main agent identify a lower amount of digital homework than those who consider the siblings, an external professional or none at all as the agent. However, for all items, the strength of association is weak as the Cramer's V value is less than .200.

TABLE II. Chi-square test of the association between the homework organisation dimension and the support agent

Category Organisation	Who helps?	Consideration		X ²	p	Cramer's V
		Low n (%)	High n (%)			
Number of homework assignments	Father	76 (33.0)	148 (67.0)	14.408	.275	-
	Mother	426 (36.7)	734 (62.3)			
	Siblings	10 (22.7)	34 (77.3)			
	Professional	14 (28.6)	35 (71.4)			
	Not required	108 (33.8)	202 (66.2)			
Difficulty of homework	Father	209 (83.3)	15 (16.7)	18.302	.107	-
	Mother	1097 (94.6)	63 (5.6)			
	Siblings	27 (61.4)	17 (38.6)			
	Professional	27 (55.1)	22 (44.9)			
	Not required	221 (71.3)	89 (28.7)			
Number of group homework assignments	Father	209 (93.3)	15 (6.7)	53.993	.000**	.100
	Mother	1097 (94.6)	63 (5.5)			
	Siblings	36 (81.8)	8 (18.2)			
	Professional	46 (93.8)	3 (6.2)			
	Not required	289 (93.3)	21 (6.8)			
Number Individual homework	Father	50 (22.3)	1074 (77.7)	10.417	.579	-
	Mother	259 (22.3)	901 (77.7)			
	Siblings	10 (22.7)	34 (77.3)			
	Professional	7 (14.3)	42 (85.7)			
	Not required	61 (19.6)	249 (80.4)			
Digital homework	Father	63 (28.1)	161 (71.9)	34.628	.001**	.080
	Mother	378 (32.6)	782 (77.4)			
	Siblings	8 (18.1)	36 (81.9)			
	Professional	11 (22.4)	38 (77.6)			
	Not required	57 (18.4)	253 (81.6)			

* = p< .050; ** p< .010

Source: Compiled by author

TABLE III. Chi-square test of the association between the homework benefit dimension and the support agent.

Category Benefits	Who helps?	Consideration		X ²	p	Cramer's V
		Low n (%)	High n (%)			
Academic	Father	118 (53.7)	106 (46.3)	10.436	.578	-
	Mother	609 (52.5)	551 (47.5)			
	siblings	20 (45.4)	24 (54.6)			
	Professional	19 (38.5)	30 (61.5)			
	Not required	147 (47.4)	163 (52.6)			
Relationship with colleagues	Father	183 (81.7)	41 (18.3)	64.990	.000**	.186
	Mother	1091 (85.4)	167 (14.6)			
	siblings	28 (63.6)	16 (36.4)			
	Professional	39 (75.5)	10 (24.5)			
	Not required	207 (66.8)	103 (33.2)			
Work habits	Father	113 (50.5)	111 (49.5)	22.730	.030*	.067
	Mother	525 (46.1)	635 (53.9)			
	siblings	20 (45.4)	24 (54.6)			
	Professional	16 (32.6)	33 (67.3)			
	Not required	138 (44.5)	172 (55.5)			
Attitude to school	Father	128 (57.1)	96 (42.9)	21.012	.057	-
	Mother	669 (57.7)	491 (42.3)			
	siblings	22 (50)	22 (50)			
	Professional	25 (51.1)	24 (48.9)			
	Not required	160 (51.6)	150 (48.4)			
Initiative and entrepreneurship	Father	130 (58.1)	94 (41.9)	65.172	.000**	.191
	Mother	708 (61.9)	452 (39.0)			
	siblings	23 (52.3)	21 (37.7)			
	Professional	24 (48.9)	25 (51.1)			
	Not required	126 (40.6)	184 (59.4)			
Family relations	Father	132 (58.9)	92 (41.1)	18.600	.099	-
	Mother	658 (56.7)	502 (43.3)			
	siblings	17 (38.7)	27 (61.4)			
	Professional	24 (48.9)	25 (51.1)			
	Not required	184 (59.4)	126 (40.6)			

* = p < .050; ** p < .010

From the Pearson Chi-Square statistic, analysing the data shown in Table III, significant differences are found according to the families' consideration of the support agent in homework with the following benefits derived from homework: improvement in the relationship with classmates ($\chi^2=54.990$, $p=.000$), development of work habits ($\chi^2 =22.730$, $p=.030$) and progress in the attitude of initiative and entrepreneurship in the academic field ($\chi^2=65.172$, $p=.000$). Analysing the frequencies in the significant associations, both in work habits, the relationship with peers and initiative and entrepreneurship, it is the families who perceive the sibling or the external professional as the main agent who have a greater identification of these benefits. With respect to Cramer's V value, more consistent relationships are found with respect to the relationship with peers and greater initiative and entrepreneurship in the academic field, given that the values are close to the moderate value (.200). Finally, it should be noted that there is no relationship of significance according to the support agent perceived by the families with respect to academic progress, positive attitude towards school or improved family relationships.

By means of the distribution analysis using Pearson's Chi-Square statistic, an important relationship of significance was found between the family's identification of the support agent for their children's homework and the children's degree of ability to cope with this educational duty, as can be seen in Table IV. In fact, the significance value is maximum ($p< .01$) in all the categories of the dimension, although, according to Cramer's V value, these differences are of greater intensity in the following skills: responsibility (Cramer's $V=.206$), time organisation (Cramer's $V=.215$) and technological ability (Cramer's $V=.234$). With the exception of the development of motivation and emotional management, in all cases, children who do not receive support are considered to have higher levels of skills. Families who indicate that they have an external professional or their own siblings as homework support are those who identify motivation and emotional management as one of the skills most present in their children when facing this task.

TABLE IV. Chi-square test of the association between the homework skills dimension and the support agent.

Skills categories	Who helps?	Consideration		X ²	p	Cramer's V
		Low n (%)	Alta n (%)			
Domain Knowledge	Father	88 (39.3)	136 (60.7)	83.092	.000**	.198
	Mother	492 (47.6)	668 (52.4)			
	Siblings	18 (40.9)	26 (59.1)			
	Professional	16 (32.7)	33 (67.3)			
	Not required	57 (17.4)	253 (82.6)			
Under-standing instructions	Father	89 (39.7)	135 (60.3)	77.051	.000**	.163
	Mother	485 (41.8)	675 (58.2)			
	Siblings	17 (38.6)	27 (61.3)			
	Professional	24 (49.0)	25 (51.0)			
	Not required	63 (20.3)	247 (79.7)			
Technologi-cal tools	Father	89 (39.7)	135 (60.3)	167.771	.000**	.234
	Mother	572 (49.3)	588 (50.7)			
	Siblings	14 (31.8)	30 (68.2)			
	Professional	12 (24.5)	37 (75.5)			
	Not required	49 (15.8)	261 (84.2)			
Time man-agement	Father	131 (59.0)	92 (41.0)	147.664	.000**	.215
	Mother	757 (65.2)	403 (34.8)			
	Siblings	14 (31.9)	30 (68.1)			
	Professional	23 (46.9)	26 (53.1)			
	Not required	105 (33.9)	205 (66.1)			
Responsi-bility	Father	112 (50.0)	112 (50.0)	135.300	.000**	.206
	Mother	656 (56.6)	504 (43.0)			
	Siblings	14 (31.9)	30 (68.1)			
	Professional	22 (44.9)	27 (55.1)			
	Not required	82 (26.4)	228 (73.6)			
Effort	Father	101 (45.1)	133 (54.9)	67.822	.000**	.192
	Mother	561 (48.3)	599 (51.7)			
	Siblings	18 (40.9)	26 (59.1)			
	Professional	20 (40.8)	29 (59.2)			
	Not required	90 (29.0)	220 (71.0)			

(Continued)

TABLE IV. Chi-square test of the association between the homework skills dimension and the support agent (Continued)

Skills categories	Who helps?	Consideration		X ²	p	Cramer's V
		Low n (%)	Alta n (%)			
Motivation	Father	140 (62.5)	84 (37.5)	70.297	.000**	.194
	Mother	738 (63.6)	422 (36.4)			
	Siblings	17 (38.6)	27 (61.4)			
	Professional	21 (49.0)	28 (57.0)			
	Not required	134 (43.2)	176 (56.8)			
Emotional management	Father	135 (60.2)	89 (39.8)	28.419	.005**	.073
	Mother	708 (61.0)	452 (39.0)			
	Siblings	17 (38.6)	27 (61.4)			
	Professional	21 (42.8)	28 (57.2)			
	Not required	167 (53.8)	143 (46.2)			

* = p<.050; ** p<.010

Source: Compiled by author

As can be seen in Table IV, as in the skills dimension, all the categories related to emotional exhaustion are significantly associated with the agent of accompaniment in school tasks perceived by families ($p=.000$). In the case of discomfort and nervousness, mental exhaustion and danger in reconciling family life, families who identify fathers and mothers as the agents are the ones who most identify these items as a detriment caused by homework accompaniment. As far as family distancing and conflict are concerned, this is more evident in families who have a professional to carry out their children's homework. Meanwhile, incapacity is more perceived as a detriment in the case of families who identify siblings as responsible for advising and monitoring their children's homework. However, the intensity of the associations results in a moderate value in the items that are more identified when parents are the agents of reference, namely: mental exhaustion (Cramer's V=.265), discomfort and nerves (Cramer's V=.201) and the imbalance of family reconciliation (Cramer's V=.202).

TABLE V. Chi-square test of the association between the dimension of emotional exhaustion caused by homework and the support agent.

Category emotional exhaustion	Who helps?	Consideration		χ^2	p	Cramer's V
		Low n (%)	Alta n (%)			
Discomfort and nervousness	Father	169 (75.2)	55 (24.8)	129.462	.000**	.201
	Mother	772 (65.5)	389 (34.5)			
	Siblings	32 (72.8)	12 (27.3)			
	Professional	37 (75.6)	12 (24.4)			
	Not required	276 (89.1)	34 (10.9)			
Mental exhaustion	Father	154 (68.8)	70 (31.2)	291.235	.000**	.265
	Mother	667 (57.5)	493 (42.5)			
	Siblings	31 (70.5)	13 (29.5)			
	Professional	39 (79.6)	10 (20.4)			
	Not required	284 (91.6)	26 (8.4)			
Disability	Father	201 (89.7)	23 (10.3)	32.745	.000**	.083
	Mother	993 (85.6)	167 (14.4)			
	Siblings	32 (72.7)	12 (27.3)			
	Professional	37 (80.5)	12 (19.5)			
	Not required	276 (89.1)	34 (10.9)			
Distance and family conflict	Father	191 (85.3)	33 (14.7)	78.525	.000**	.196
	Mother	896 (77.2)	264 (22.8)			
	Siblings	36 (81.8)	8 (18.2)			
	Professional	36 (73.5)	13 (26.5)			
	Not required	282 (90.9)	28 (9.1)			
Attempts to reconcile work and family life	Father	145 (64.7)	79 (35.3)	88.917	.000**	.202
	Mother	701 (60.4)	459 (39.6)			
	Siblings	33 (75.0)	11 (25.0)			
	Professional	34 (69.4)	15 (30.6)			
	Not required	254 (72.0)	56 (28.0)			

* = p<.050; ** p<.010

Source: Compiled by author

Discussion and conclusions

Traditionally, the education and upbringing of offspring has been a function biologically and culturally attributed to women (Hernández-Prados et al., 2020). After years of fighting for equality, and the proliferation of pedagogical movements calling for responsible, involved

and committed parenting (Rodrigo et al., 2015), this study shows that the person who usually helps with homework is mostly the mother, followed far behind by the father, and to a lesser extent by siblings. It should be remembered that the study is based on families' perceptions of homework, which indicate in each case who is the main agent providing help and how they perceive each of the dimensions analysed. In this way, we understand that it is women who are primarily involved in helping with their children's educational tasks, as supported by most studies in this field (Fernández-Freire et al., 2019; Gónida and Cortina, 2014; Hernández-Prados and Gil-Noguera, 2022; Valdés-Cuervo et al., 2022). In this sense, Lehner-Mear's (2021) analysis of the gender impact of homework in families attributes this differentiation to cultural pressure, as it is perceived that the label of good mother is given to those who help with homework, and conversely, bad mothers to those who resort to other alternative practices such as external support. In short, homework support is conceived by mothers as an unequivocal act of love for others (Páez and Zúñiga, 2021).

On the other hand, the research carried out reveals, in general, significant differences in relation to who helps with homework (fathers, mothers, brothers and sisters, external professional or doing homework alone) and the dimensions of organisation of homework, perceived benefits, development of skills and emotional exhaustion produced in the family. Thus, with regard to the task organisation variable, corresponding to the first objective, we can conclude that, regardless of who accompanies them, families value individual tasks much more highly than group tasks. Perhaps because the teacher depends too much on individual activities programmed in textbooks (Pineda and Fraile, 2020) and group activities are sporadic (Drechsler, 2021; Kaur, 2011). Only families in which siblings help out consider group work to be significantly important, probably because the generation gap is smaller. Regarding digital homework, it is significantly more highly valued in families that delegate professionals or siblings, while in families where parents, especially mothers, help out, there is more reticence towards this type of homework. This digital divide has been widely studied, despite the fact that "new technologies and group work seem to function as an "escape valve"" (Pineda and Fraile, 2020, p.298).

In general, in relation to the second objective, the data show that families' perception of homework is positive, considering that it provides

multiple benefits (Orozco-Vargas et al., 2022; Pomerantz et al., 2007). Homework favours, on the one hand, the acquisition of work habits, as has already been shown in other studies Tan et al. (2020), although significantly in families who use private academies, and on the other hand, initiative and entrepreneurship, especially when homework is done independently, as recognised by Feng et al. (2019). In contrast, homework does not contribute to fostering peer relationships and a favourable attitude towards school, especially during the isolation period (Drechsler, 2021), although families with sibling helpers are the ones who value these attribution benefits the most.

However, this issue is not without controversy and, as if it were two sides of the same coin, academic homework is also perceived negatively, altering intrafamily relationships and emotional health (Sánchez-Lissen, 2015), the book *Cómo sobrevivir a los deberes de tu hijo* by Bailén (2016) is an example of this. More specifically, the data indicate, in relation to the third objective, that mothers stand out significantly in conceiving homework as something that makes them uncomfortable, stresses them, exhausts them mentally and makes it difficult to reconcile family life, showing a tendency contrary to that found in previous studies that recognise that parents find this role of helping unpleasant and stressful (Páez and Zúñiga, 2021). Sibling apprenticeships are sometimes used to free parents and facilitate reconciliation. However, the inability to assume support in school duties is the main constraint significantly attributed to siblings, although Reimer and Colina-Hill (2022) noted emotional problems and burnout. Finally, the performance of homework is a conflictive issue that deteriorates family relationships (Bai-lén and Polo, 2016), preferring to delegate this function to external professionals. Hence, this difficulty has been significantly highlighted by families who opt for these alternative practices.

Finally, the families participating in the study attributed significantly more skills to children who tackle homework autonomously, without help or mediation, except for the ability to organise time, motivation and emotional management, which are assigned to children who receive help from their siblings. Autonomy with respect to homework is the goal desired by teachers and families, as it denotes self-efficacy in various capacities (Feng et al., 2019; Hernández-Prados and Gil-Noguera, 2022; Valdés-Cuervo et al., 2022). Moreover, collaboration with siblings enables the development of the more emotional level (Van der Kaap-Deeder et al., 2017).

Research is not without limitations, most of which are the result of the choices made. In the present study, some of them have been identified. Firstly, with regard to the coverage of the data, being a cross-sectional study, it does not allow us to check how the way of coping with homework evolves as autonomy is achieved and does not require family accompaniment. Secondly, data could have been collected from adolescents in order to be able to compare whether there are differences with respect to the perception of parents.

From the position in the emerging line of research that questions family support and its relationship to academic success, focusing on how this support is given rather than how much (Fernández-Alonso et al., 2022), the present study adds new elements of discussion that contribute to scientific, academic and family development. Specifically, some questions such as: What type of help is provided, what impact does accompaniment have on the children and on the helper, how does the helper value homework, what are the reasons for doing so, why do men delegate this family responsibility to women, require further analysis, since, as Valle et al. (2015) point out, it is no longer a question of the amount of time invested in doing homework, but rather of the use of time and efficiency in accompaniment.

Finally, seeking and promoting aspects that guarantee adequate support for each of the agents involved is an essential element in improving the educational reality, compensating for difficulties and making it possible to get more out of homework in terms of the benefits and competences attributed to it. In this respect, the results obtained contribute to the understanding that homework from the perspective of co-responsibility implies that its improvement does not depend exclusively on family members. Teachers should receive specific training on homework to prevent it from becoming repetitive (Feito-Alonso, 2020), and to promote quality tasks that encourage debate, exchange, the establishment of rules and autonomous responsibility (Álvarez-Blanco, 2019). All of this will result in a healthier family climate.

Bibliographical references

Álvarez-Blanco, L. (2019). Modelos teóricos de implicación educativa familiar: responsabilidades compartidas entre centros educativos, familias y comunidad. *Aula Abierta*, 48(1), 19-30. <https://doi.org/10.17811/rifie.48.1.2019.19-30>

Arias-Redondo, M., Latas-Pérez, C., & García-Gómez, A. (2006). Los deberes a examen. *Aula de innovación educativa*, 153-154, 46-51. <http://hdl.handle.net/11162/85875>

Bailén, E. (2016). *Cómo sobrevivir a los deberes de tu hijo*. Ediciones Martínez Roca.

Bailén, E., & Polo, I. (2016). Deberes escolares: el reflejo de un sistema educativo. *Avances en Supervisión Educativa*, 25, 1-36. <https://doi.org/10.23824/ase.v0i25.543>

Bempechat, J. (2019). The case for (quality) homework: Why it improves learning, and how parents can help. *Education Next*, 19(1), 36-44. <https://www.educationnext.org/case-for-quality-homework-improves-learning-how-parents-can-help/>

Drechsler, A. (2021). From synchronous face-to-face group work to asynchronous individual work: Pivoting an enterprise modeling course for teaching during a COVID-19 lockdown. *Communications of the Association for Information Systems*, 48(1), 22. <https://doi.org/10.17705/1CAIS.04822>

Epstein, J. L. (1990). *School, family and community partnership*. Corwin Press, INC.

Feito-Alonso, R. (2020). Los deberes escolares. Un análisis sistematizado con especial referencia al caso español. *Contextos Educativos. Revista de Educación*, 25, 163-182. <https://doi.org/10.18172/con.3957>

Feng, X., Xie, K., Gong, S., Gao, L., & Cao, Y. (2019). Effects of Parental Autonomy Support and Teacher Support on Middle School Students' Homework Effort: Homework Autonomous Motivation as Mediator. *Frontiers Psychology*, 10, 1-11. <https://doi.org/10.3389/fpsyg.2019.00612>

Fernández-Alonso, R., Álvarez-Díaz, M., García-Crespo, F. J., Woitschach, P., & Muñiz, J. (2022). Should We Help our Children with Homework? A Meta-Analysis Using PISA Data. *Psicothema*, 34(1), 56-85. <https://doi.org/10.7334/psicothema2021.65>

Fernández-Freire, L., Rodríguez, B., & Martínez, R. A. (2019). Padres y madres ante las tareas escolares: La visión del profesorado. *Aula Abierta*, 48(1), 77-84. <https://doi.org/10.17811/rifie.48.1.2019.77-84>

García-Sanz, M. P., Gomariz-Vicente, M. A., Hernández-Prados, M. A., & Parra-Martínez, J. (2010). La comunicación entre la familia y el centro educativo, desde la percepción de los padres y madres de los alumnos. *Educatio Siglo XXI*, 28(1), 157-188 <https://revistas.um.es/educatio/article/view/109771>

Gónida, E. N., & Cortina, K. S. (2014). Parental involvement in homework: Relations with parent and student achievement-related motivational beliefs and achievement. *British Journal of Educational Psychology*, 84(3), 376-396. <https://doi.org/10.1111/bjep.12039>

Grolnick, W. S., Benjet, C., Kurowski, C. O., & Apostoleris, N. H. (1997). Predictors of Parent involvement in Children's Schooling. *Journal of Educational Psychology*, 89(3), 538-548. <https://doi.org/10.1037/0022-0663.89.3.538>

Hernández-Prados, M. A., Belmonte, M., & Lara Guillén, B. (2020). El reflejo de una tradición: Feminización de la vida familiar. *ReiDoCrea: Revista Electrónica de Investigación y Docencia Creativa*, 9, 232-243. <http://dx.doi.org/10.30827/Digibug.66364>

Hernández-Prados, M. A., & Gil-Noguera, J. A. (2018). Los deberes a examen. Riesgos y oportunidades. En J. C. Martínez Coll (Ed.), *III Congreso online internacional sobre la educación del siglo XXI* (pp. 116-129). Eumed. <https://www.eumed.net/actas/18/educacion/9-los-deberes-a-examen-riesgos.pdf>

Hernández-Prados, M. A., & Gil-Noguera, J. A. (2022). El papel de la familia en la realización de los deberes escolares. *Revista Electrónica Educare*, 26(2), 1-18. <https://doi.org/10.15359/ree.26-2.16>

Hoover-Dempsey, K., & Sandler, H. M. (1995). Parental involvement in children 's education: why does it make a difference? *Teachers College Record*, 97(2), 310-331. https://www.academia.edu/1338702/Parental_involvement_in_childrens_education_Why_does_it_make_a_difference

Kaur, B. (2011). Mathematics homework: A study of three grade eight classrooms in Singapore. *International Journal of Science and Mathematics Education*, 9(1), 187-206. <https://doi.org/10.1007/s10763-010-9237-0>

Kohn, A. (2013). *El mito de los deberes: ¿Por qué son perjudiciales para el aprendizaje y la convivencia?* Kaleida.

Lehner-Mear, R. (2021). Good mother, bad mother?: Maternal identities and cyber-agency in the primary school homework debate. *Gender and Education*, 33(3), 285-305. <https://doi.org/10.1080/09540253.2020.1763920>

Martínez-Vicente, M., Suárez-Riveiro, J. M., & Valiente-Barroso, C. (2020). Implicación estudiantil y parental en los deberes escolares: diferencias según el curso, género y rendimiento académico. *Revista de Psicología y Educación*, 15(2), 151-165. <http://dx.doi.org/10.23923/rpye2020.02.193>

Murillo, F. J., & Martínez-Garrido, C. (2013). Incidencia de las tareas para casa en el rendimiento académico. Un estudio con estudiantes iberoamericanos de Educación Primaria. *Revista de Psicodidáctica*, 18(1), 157-178. <https://doi.org/10.1387/RevPsicodidact.6156>

Mylonakou, I., & Kekes, I. (2005). Syneducation (Synekpaidefsis): *Reinforcing Communication and Strengthening Cooperation among Students, Parents and Schools*. Cambridge. <http://www.hfrp.org/content/download/1262/48765/file/syneducation.pdf>

Núñez, J. C., Epstein, J. L., Suárez, N., Rosário, P., Vallejo, G., & Valle, A. (2017). How Do Student Prior Achievement and Homework Behaviors Relate to Perceived Parental Involvement in Homework? *Frontiers in Psychology*, 8, 1-13. <https://doi.org/10.3389/fpsyg.2017.01217>

Olson, D., & Gorall, D. (2006). *FACES IV&the Circumplex Model*. Life Innovations.

Orozco-Vargas, A. E., Aguilera-Reyes, U., García-López, G. I., & Venebra-Muñoz, A. (2022). Funcionamiento Familiar y Autoeficacia Académica. *Revista de Educación*, 396, 127-150. <https://doi.org/10.4438/1988-592X-RE-2022-396-532>

Páez, I., & Zúñiga, M. (2021). ¿Paternidades en transformación? Ser padre en Culiacán, Sinaloa, en tiempos de confinamiento y crisis sanitaria. *Región y sociedad*, 33, 1-23. <https://doi.org/10.22198/rys2021/33/1502>

Pérez, R. A. P., Reyes, D. A. L., Urquico, L. J. L., & Vicente, H. C. D. (2020). Students' Perception on the Use of Homework for Learning. *Ascendens Asia Journal of Multidisciplinary Research Abstracts*, 4(1). <https://ojs.aaresearchindex.com/index.php/AAJMRA/article/view/2708>

Pineda, J. A., & Fraile, F. J. (2020). El modelo didáctico como articulador del sistema-aula: un estudio de caso en educación secundaria.

Estudios pedagógicos, 46(1), 285-300. <https://doi.org/10.4067/s0718-07052020000100285>

Pomerantz, E. M., Moorman, E. A., & Litwack, S. D. (2007). The How, Whom, and Why of Parents' Involvement in Children's Academic Lives: More Is Not Always Better. *Review of Educational Research*, 77(3), 373-410. <https://doi.org/10.3102/003465430305567>

Pressman, R., Sugarman, D., Nemon, M. L., Desjarlais, J., Owens, J. A., & Schettini-Evans, A. (2015). Homework and Family Stress: With Consideration of Parents' Self Confidence, Educational Level, and Cultural Background. *The American Journal of Family Therapy*, 43(4), 297-313. <https://doi.org/10.1080/01926187.2015.1061407>

Reimer, T., & Hill, J. C. (2022). Crossing the digital divide and the equity expanse: reaching and teaching all students during the pandemic. *Journal of Leadership, Equity, and Research*, 8(1), 71-86. <https://journals.sfu.ca/cvj/index.php/cvj/issue/view/25/75>

Rodrigo, M^a. J., Máiquez, M. L., Martín, J. C., Byrne, S., & Rodríguez, B. (Eds.). (2015). *Manual práctico de parentalidad positiva*. Síntesis

Rodríguez-Ruiz, B., Martínez-González, R. A., & Ceballos-Vacas, E. M. (2019). Las tutorías con las familias en Educación Secundaria Obligatoria: Percepción de padres, madres y profesorado tutor. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 22(3), 31-43. <https://doi.org/10.6018/reifop.389351>

Rosário, P., Carlos Núñez, J., Vallejo, G., Nunes, T., Cunha, J., Fuentes, S., & Valle, A. (2018). Homework purposes, homework behaviors, and academic achievement: Examining the mediating role of students' perceived homework quality. *Contemporary Educational Psychology*, 53, 1-53. doi: <https://doi.org/10.1016/j.cedpsych.2018.04.001>

Sáez-López, M. J. (2017). *Investigación educativa. fundamentos teóricos, procesos y elementos prácticos*. Editorial UNED.

Sallee, B., & Rigler, N. (2008). Doing our homework on homework: How does homework help? *English Journal*, 98(2), 46-51. <http://www.jstor.org/stable/40503382>

Sánchez Lissen, E. (2015). Los deberes escolares en casa. *Diálogo familia colegio*, 309, 29-38.

Tan, C. Y., Lyu, M., & Peng, B. (2020). Academic benefits from parental involvement are stratified by parental socioeconomic status: A meta-analysis. *Parenting*, 20(4), 241-287. <https://doi.org/10.1080/15295192.2019.1694836>

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology*, 98, 438-456. <https://doi.org/10.1037/0022-0663.98.2.438>

Valdés-Cuervo, A. A., Grijalva-Quiñonez, C. S., & Parra-Pérez, L. G. (2022). Apoyo parental a la autonomía y terminación de las tareas: Efectos mediadores de la autoeficacia académica, el propósito y las emociones de los niños al realizar las tareas. *Anales de Psicología*, 38(2), 259-268. <https://doi.org/10.6018/analesps.424221>

Valiente-Barroso, C., Suárez-Riveiro, J. M., & Martínez-Vicente, M. (2020). Implicación estudiantil y parental en los deberes escolares: diferencias según el curso, género y rendimiento académico. *Revista de Psicología y Educación*, 15(2), 151-165. <https://doi.org/10.23923/rpye2020.02.193>

Valle, A., Núñez, J. C., Cabanach, R. G., Rodríguez, S., Rosário, P., & Inglés, C. (2015). Motivational profiles as a combination of academic goals in higher education. *Educational Psychology*, 35(5), 634-650. <https://doi.org/10.1080/01443410.2013.819072>

Valle, A., Rodríguez-Prado, P., Regueiro, B., Estévez, I., Piñeiro, I., & Rodríguez, S. (2021). Un estudio comparativo de los deberes escolares en el alumnado extranjero y no extranjero. *Estudios pedagógicos*, 47(2), 53-78. <https://doi.org/10.4067/S0718-07052021000200053>

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>

Wilder, S. (2014). Effects of parental involvement on academic achievement: a meta-synthesis. *Educational Review*, 66(3), 377-397. <https://doi.org/10.1080/00131911.2013.780009>

Contact address: Juan Antonio Gil Noguera, Universidad de Murcia, Facultad de Educación y Departamento de Teoría e Historia de la Educación. Campus de Espinardo, 12, 30100 Murcia, Spain. E-mail: juanantonio.gil@um.es

The paradox of learning to learn development: adapting a scale in secondary school students

La paradoja del desarrollo de aprender a aprender: adaptación de una escala en estudiantes de secundaria

<https://doi.org/10.4438/1988-592X-RE-2023-401-582>

José Francisco Martín Alonso

<https://orcid.org/0000-0003-0479-5771>

Universidad Pontificia Comillas y Universidad Internacional de la Rioja

Isabel Muñoz San Roque

<https://orcid.org/0000-0003-1757-5700>

Universidad Pontificia Comillas de Madrid

Abstract

The conceptualization of the competence of learning to learn has evolved over the years, and its name has been changed to personal, social, and learning to learn competence. However, the analysis of this competence's cognitive and metacognitive development is still central when talking about the self-regulation of learning, especially at the secondary school stage where the work of guidance departments is fundamental in developing specific skills for learning. This article presents a twofold objective: adapting a scale to measure the self-perception of the level of development of this competence, focusing on the cognitive and metacognitive dimension in secondary school students, and, secondly, the analysis of its evolution by year. The sample comprises 1033 secondary school students from the Community of Madrid. The first objective involves a double exploratory and confirmatory factor analysis process, selecting two different samples. The criterion validity is based on the relationship of competence with learning approaches and academic self-efficacy. An ANOVA and a Student's t-test were performed for the second objective. The results show adequate internal

consistency, obtaining a factorial structure of three factors (Self-assessment of the process, Self-knowledge as a learner, and Management of the learning process) and goodness of fit indices adequate to the postulated theoretical model, in addition to the relationships to support criterion validity are coherent and significant ($p < .05$). The results indicate that the level of development of learning to learn competence decreases as the academic year's progress. There are emotional aspects, such as motivation, which could explain this decline throughout secondary education, and it is essential to be able to act on them in the guidance processes.

Keywords: learning processes, competence, self-regulation, metacognition, self-efficacy, learning motivation, secondary education.

Resumen

La conceptualización de la competencia de aprender a aprender ha ido evolucionando a lo largo de los años, llegando a modificarse su denominación por competencia personal, social y de aprender a aprender. Sin embargo, el análisis del desarrollo cognitivo y metacognitivo de esta competencia sigue siendo central al hablar de la autorregulación del aprendizaje, especialmente en la etapa de secundaria donde la labor de los departamentos de orientación es fundamental en el desarrollo de ciertas habilidades para aprender mejor. Este artículo presenta un doble objetivo, el proceso de adaptación de una escala para medir la autopercepción del nivel de desarrollo de esta competencia, centrándose en la dimensión cognitiva y metacognitiva, en alumnos de secundaria y, en segundo lugar, el análisis de su evolución por curso. La muestra está formada por 1033 estudiantes de educación secundaria de la Comunidad de Madrid. El primer objetivo supone un doble proceso de análisis factorial exploratorio y confirmatorio seleccionando dos muestras diferentes. La validez criterial se apoya en la relación de la competencia con los enfoques de aprendizaje y la autoeficacia académica. Para el segundo objetivo se realiza un ANOVA y una t de Student. Los resultados muestran una adecuada consistencia interna, obteniendo una estructura factorial de tres factores (Autoevaluación del proceso, Autoconocimiento como aprendiz y Gestión del proceso de aprendizaje) y unos índices de bondad de ajuste adecuados al modelo teórico postulado, además las relaciones para apoyar la validez criterial son coherentes y estadísticamente significativas ($p < .05$). Los resultados indican que el nivel de desarrollo de la competencia de aprender a aprender disminuye a medida que avanzan los cursos académicos. Hay aspectos emocionales, como la motivación, que podrían explicar este descenso a lo largo de la educación secundaria y es importante poder actuar sobre ellos en los procesos de orientación.

Palabras clave: procesos de aprendizaje, competencia, autorregulación, metacognición, autoeficacia, motivación para el aprendizaje, educación secundaria.

Introduction

The theoretical approach used to define the learning to learn competence (also known as the LTL competence) is that developed by the European conceptualisation, research and development framework (Caena & Punie, 2019; European Commission, 2006; European Council, 2018; Fredriksson & Hoskins, 2006a; García et al., 2022; Hoskins & Fredriksson, 2008; Hutmacher, 1997; Sala et al., 2020; Salas & Gallardo, 2022; Stringher, 2014; Stringher et al., 2021; Valle, 2020), which has been guiding the state legislation governing this competence (LOE, 2006; LOMCE, 2013; LOMLOE, 2020; Martín-Alonso, & Muñoz-San Roque, 2022) with a perspective centred on the concept of self-regulatory competence (Hadwin et al., 2018; Salas & Gallardo, 2022; Salmerón et al., 2010; Schunk & Zimmerman, 1997; Torre, 2007; Usher & Schunk, 2018; Zimmerman 1995). Thus, the learning to learn competence is understood as the set of cognitive, metacognitive, emotional and relational skills that allow students to be aware of and to manage their own learning.

The two major institutional developments of the learning to learn competence in the EU took place at the two main moments in which the EU selected and defined the key competences (European Commission, 2006; European Council, 2018) within its strategies for the years 2010 and 2020. As part of this process of conceptualisation, research and development, there has been an evolution leading to a change in the terminology used, from the old term – learning to learn competence (European Commission, 2006) – to the new term – personal, social and learning to learn competence (European Council, 2018).

Since its beginnings in the European context, two perspectives of this competence were noted: the cognitive psychology paradigm and the sociocultural paradigm, exemplified by the figures of Piaget and Vygotsky (Hoskins & Fredriksson, 2008; Fredriksson & Hoskins, 2006a). Hence, in the general framework of the Education and Training 2010 project (European Council, 2000), a series of meetings of experts on the subject were organised and they convened the Learning to Learn Network meetings (Fredriksson & Hoskins, 2006a; 2006b; 2007), focusing their work on the conceptualisation and measurement of the learning to learn competence. The aim was to create an instrument that was subsequently applied in eight European Union countries (Kupiainen et al., 2008).

The inconclusive results obtained through the instrument and the withdrawal of funds by the European Union resulted in a dearth of publications on the subject in the Centre for Research on Lifelong Learning (CRELL) until the issue was taken up again in the new 2018 reference framework of competences (Martín-Alonso, 2021). The new definition of the learning to learn competence (European Council, 2018) has prompted a new conceptualisation, research and development framework continuing on from the previous framework (Caena & Punie, 2019; Caena & Stringher, 2020; García et al., 2022; Sala et al., 2020; Suárez et al., 2005; Valle, 2020). The concept continues to evolve in a broader context around metacognition and self-regulation of the traditional definition and introducing socio-emotional aspects, well-being and health (Caena & Punie, 2019). Thus, the current definition contemplates aspects such as collaboration with others, the contribution to physical and emotional well-being, healthy living and conflict management, without rejecting existing aspects, such as time management and learning (European Council, 2018).

In this context of the evolving thinking on the concept of learning to learn, not only is this understood from the perspective of self-regulatory competence (Hadwin et al., 2018; Salmerón et al., 2010; Schunk & Zimmerman, 1997; Torre, 2007; Usher & Schunk, 2018; Zimmerman, 1995) and as a broad concept that encompasses metacognition (Efklides, 2011; Moreno & Martín, 2007; Panadero & Tapia, 2014; Pintrich, 2000; Torre, 2007; Whitebread et al., 2007; Winne 2018; Zimmerman, 1995), but it also integrates cognitive and sociocultural perspectives of learning (García et al., 2022; Hadwin et al., 2018; Panadero, 2017; Panadero & Tapia, 2014; Salas & Gallardo, 2022; Schunk & Greene, 2018; Usher & Schunk, 2018; Winne, 2018; Zimmerman, 2013).

Thus, based on the analysis of the aforementioned theoretical models, we have selected the most significant dimensions that determine the concept of the learning to learn competence (Martín-Alonso, 2021; Muñoz-San Roque et al., 2016): self-assessment of the process; self-awareness as a learner; and management of the learning process. These dimensions differ from the current theoretical framework for defining the personal, social and learning to learn competence, as they omit the more social part of learning, which has been included in the latest legislation. Further research will add this aspect when designing and assessing this competence, an aspect that is considered fundamental in the new European conceptualisation framework (European Council, 2018) and which

has inspired a change in the legislation in Spain (LOMLOE, 2020). The instrument adapted in this study has focused on the cognitive aspects of the competence, following the legislative framework of the moment. However, a major contribution is that it addresses personal aspects of the student, including self-awareness as a learner.

The first of the three dimensions considered relevant when defining the learning to learn competence from cognitive and metacognitive perspectives refers to the dimension of self-assessment of the process, that is, the verification of the procedure followed by the student while learning. It is a basic metacognitive strategy within the learning to learn competence (Hautamäki et al., 2002; Zimmerman, 2013) that, in addition to knowledge of the process, assumes executive control during learning that incorporates control of the effort involved in the task (Martínez-Fernández, 2007).

The second dimension focuses on self-awareness as a learner, which is what Deakin-Crik (2014) call strategic knowledge, that is, being aware of one's own learning process and of the context by applying metacognitive strategies such as being aware of oneself (cognitively and affectively), of the learning process and of the relationship between oneself and that process (Villardón-Gallego et al., 2013). It has an emotional aspect, as it integrates the awareness of strengths and weaknesses and, in general, emotional self-awareness (Martín & Moreno, 2007).

Thirdly, management of the learning process includes the planning and the cognitive management of the process, whereby the student sets realistic goals and implements metacognitive strategies in order to acquire a personal commitment to achieving these goals (Martínez-Fernández, 2007; Villardón-Gallego et al., 2013; Zimmerman, 2013). It incorporates aspects such as planning (times, schedules, etc.), cognitive management and testing and, ultimately, knowledge and control of strategies adapted to the task (Caena & Punie, 2019).

When analysing how the learning to learn competence evolves as students progress through academic years, there are several studies that show that its development decreases, as do motivational aspects (Gaeta, 2013; González Fernández, 2005; Palomo del Blanco, 2014; Rodríguez Fuentes, 2009; Rosario et al., 2012). In this regard, Dignath and Büttner (2008) conducted a meta-analysis of self-regulation in primary and secondary school and concluded that students in higher years learn more strategically, but their motivation decreases as they progress from year to year. Stringher (2021) considers that, in the development of learning

competences, there are key elements such as motivation, creativity and curiosity to learn and that these decrease over the school years, to increase again in early adulthood. This decline in the development of learning skills due to motivational factors also appears in the classic study by Zimmerman and Martínez-Pons (1990) and is also noted by Pintrich (2003) with reference to the decline in student motivation.

This paper presents the most relevant results of this research, the main objectives of which are twofold.

The first objective is the adaptation of a valid, reliable instrument that measures the development of the learning to learn competence in secondary school students, taking into account the legislative framework that defined the cognitive and metacognitive aspects of the competence.

The second objective is to analyse whether the paradox of a decrease in their perception of their development of the competence is present as students progress to higher years. This is an issue of interest in the field of educational research and for the work carried out by professionals in psycho-pedagogical guidance, since the development of learning skills has always been a fundamental field of work in schools and the object of tutorial action.

Method

The research design was cross-sectional and used quantitative methodology. The approach applied in the process of adapting the scale was hypothetical-deductive, based firstly on theoretical sources and then, through empirical validation, based on a structure with different factors from a scale validated in university students.

Sample

The sample was a non-probabilistic incidental sample and met three fundamental requirements: the subjects were secondary school students from state, subsidised or private schools in the Region of Madrid, they were studying the different academic options available in the fourth year of Obligatory Secondary Education (ESO) and they were enrolled in different years. During the 2017-2018 and 2018-2019 academic years, the questionnaire was

given to 1155 subjects. Missing or randomly answered values were eliminated. Data were selected from a sample of 1033 secondary education and baccalaureate students, of whom 544 were female and 489 were male. 163 (15.8%) were from the first year of ESO (1ESO), 184 (17.8%) from the second year of ESO (2ESO), 200 (19.4%) from the third year of ESO (3ESO), 336 (32.5%) from the fourth year of ESO (4ESO), 130 (12.6%) from the first year of baccalaureate and, finally, 18 (1.7%) from the second year of baccalaureate. As regards the schools, 759 students studied in subsidised schools (72.7%), 136 in private schools (13.2%) and 146 in state schools (14.1%).

For the dual process of confirmatory and exploratory analysis, the sample was divided into two according to the dates on which the questionnaire was answered. The sample for the exploratory analysis consisted of 355 participants, of whom 219 (61.7%) studied in subsidised schools and 136 (38.3%) in private schools. The sample selected for the confirmatory factor analysis (CFA) consisted of 678 participants, of whom 530 (78.4%) studied at subsidised schools and 146 (21.6%) at state schools. With respect to the number of subjects required to carry out a CFA, Ferrando and Anguiano-Carrasco (2010) suggest a sample of 200 observations as a minimum to be taken into account. Rojas-Torres (2020) also advocates samples of 200 for a CFA, indicating that from this N onwards the increase does not greatly affect the robustness of the classical adjustment indices.

Instruments

The objective of adapting a scale to measure the learning to learn competence in secondary school students involved an initial analysis of existing instruments and the detailed selection and study of those that were considered essential. The starting point was an instrument that had already been developed and validated, aimed at university samples (Muñoz-San Roque et al., 2016). It was adapted to adequately take into account the differential characteristics of a sample of secondary school students. The analysis began with a construct of self-regulatory competence (Hadwin et al., 2018; Salmerón et al., 2010; Schunk & Zimmerman, 1997; Usher & Schunk, 2018; Zimmerman, 1995), with the integrating vision of Torre (2007). The basis, then, was a structure comprised of four components (context, cognition, behaviour and emotion) with two subdimensions in each component (knowledge and control/management). Similarly, the

process of elaboration of the European pre-pilot test to measure the learning to learn competence (Hoskins & Fredriksson, 2008) was taken into consideration, as were the existing instruments (Deakin-Crick et al., 2004; Elshout-Mohr et al., 2004; Hautamäki et al., 2002; Moreno, 2002), which led, in 2008, to the pre-pilot test applied in eight countries (Kupiainen et al., 2008; Moreno et al., 2008).

The first stage was the drafting of 26 items. For the selection of the definitive items, two fundamental criteria were applied: content validation by university professors through the Clarity, Appropriateness, Relevance and Accessibility (CARA) model (Hernández-Franco & Gonzalo-Misol, 2009) and psychometric analysis of the operation items, maintaining those whose highest factor saturation in the rotated matrix fell within the dimension postulated by the proposed theoretical model and which did not have a weighting greater than .30 in the rest of the factors.

After the analyses, 19 items were maintained on a scale of 1 to 6 (not very developed to highly developed) and which maintained the different conceptual nuances on which the theoretical basis rested.

The analyses also included:

- Sociodemographic data (year, school, sex, age, number of curricular subjects failed and average grade in the last evaluation, academic self-perception, academic options taken in the fourth year of ESO being studied, academic option at the end of the fourth year of ESO, parents' level of studies)

Two instruments were also applied in order to analyse the criterion validity of the Scale:

- The Revised Two-Factor Study Process Questionnaire (R-LPQ-2F) scale (Kember et al., 2004), translated by González Geraldo et al. (2010).
- The Academic Self-Efficacy Scale (Torre, 2007).

Procedure and data analysis

The battery of tests was designed in paper format and those responsible for the schools, who collaborated voluntarily, were asked to administer

the questionnaires. The internal protocols applied in the schools guaranteed that the questionnaires were answered confidentially and voluntarily and that the required consent was given, thus guaranteeing the ethical criteria of the data gathering process.

The internal consistency coefficients were calculated using Cronbach's Alpha and McDonald's Omega coefficient, since the scale of the items was ordinal and there were fewer than seven response options (McDonald, 1999). The homogeneity indices were calculated using the IBM SPSS Statistics 20.0 and Jamovi statistical packages. Similarly, construct validity was analysed by means of an exploratory factor analysis (principal component analysis and Promax rotation), certifying the sample adequacy of the scale by means of the Kaiser-Meyer-Olkin test and Bartlett's sphericity test.

Criterion validity was checked by analysing the correlation of the scale and its factors with variables, namely the deep focus, shallow focus and self-efficacy variables. These are constructs that have been established in educational research and with which the scientific literature indicates a relationship.

To perform the confirmatory factor analysis, structural covariance techniques were used since, according to Martínez-Abad and Rodríguez-Conde (2017), the estimates of the product-moment or polychoric correlation coefficients are very similar when the number of response options is greater than 5 in ordinal variables. For this purpose, the EQS 6.1 (Structural Equation Modelling Software) program (Bentler, 1995) was used. For parameter estimates, the robust maximum likelihood (RML) method was used, with a series of indicators to assess the model fit, in accordance with various authors (Abad et al., 2011; Brown, 2006; Byrne, 2006; Cho et al., 2020; De Carvalho & Chima, 2014; Fan et al., 2016; Goh & Yusuf, 2017; González-Montesinos & Backhoff, 2010; Hair et al., 1998; Hu & Bentler, 1999; Jöreskog, 1970; Kline, 2005; Xia & Yang, 2019). Thus, absolute χ^2 and relative χ^2 fit indicators (divided by degrees of freedom), Akaike's Information Criterion (AIC), the root mean square error of approximation (RMSEA) index, GFI (goodness of fit index), CFI (confirmatory fit index), TLI (Tucker and Lewis index) and the standardised root mean square residual (SRMR) index were used. The final model will be presented with the parameters of the structural relationships indicating the standardised factor coefficients and the estimation errors.

To test the second research objective, the Students' t statistics and analysis of variance (factorial ANOVA) were used to assess the differences between groups. Information on the Students' t or the F in ANOVA, the probability that the difference is due to chance (p), the degrees of freedom (df) and the effect size (d or η^2) will be shown. An analysis was performed of the assumptions of homogeneity of variances, using Levene, and of assumptions of normality, using Shapiro-Wilk. Nonparametric tests (Mann-Whitney U and Kruskal-Wallis) were analysed when they were not met in order to confirm that the results held. Significant values were considered to be those where $p<.05$.

Results

In response to the first research objective, to assess the internal consistency of the scale, values of .888 in Cronbach's Alpha and .891 in McDonald's Omega coefficient were obtained for the 19-item scale, which allows us to conclude that the scale adequately discriminates between students in their perception of their development of the learning to learn competence. An analysis of the homogeneity indices of the items (correlation of each item with the total scale without the item) shows that all have values above .32. Table I shows the means and standard deviations of the items and dimensions, the homogeneity indices of the items, and the Cronbach's Alpha and McDonald's Omega coefficient of the scale and its dimensions.

TABLE I. Descriptive statistics and internal consistency analysis

Total and factors	Mean	Deviation	Alpha	Omega
Perceived level of development of the Learning to Learn competence (Total)	4.20	0.77	.888	.891
Self-awareness as a learner	4.52	0.78	.723	.726
Self-assessment of the process	4.20	0.90	.770	.780
Learning management	3.86	1.02	.767	.772

(Continued)

TABLE I. Descriptive statistics and internal consistency analysis (Continued)

ITEMS IN THE SCALE	Mean	Devi-a-tion	r	Omega without item
Self-awareness as a learner				
8. I use different ways of studying depending on the task I am asked to do.	4.19	1.450	.445	.692
10. I am aware of my strengths and weaknesses when I am studying or learning (I know what I am good or bad at).	5.08	1.161	.375	.704
11. I ask for help from the right person when I need it.	4.58	1.417	.450	.690
12. I like learning.	4.29	1.383	.435	.691
13. I feel capable of successfully completing the learning tasks in order to achieve the objectives proposed in the subjects.	4.43	1.201	.508	.673
14. When I am learning I think of content related to other subjects or things I already know.	4.18	1.290	.441	.692
2. I am aware of the value of learning for the people around me (parents, teachers, etc.).	4.85	1.161	.346	.714
Self-assessment of the process				
5. I check whether I am doing well in studying for an exam or doing a learning task or if I need to change the way I do it.	4.06	1.158	.567	.735
6. I know the steps I am taking when I am studying and I can describe them orally.	4.23	1.336	.533	.743
9. I make the effort necessary to learn.	4.34	1.273	.541	.740
17. I am sufficiently proficient in study techniques such as underlining, outlining, summarising, etc.	4.15	1.534	.428	.771
18. I have good, effective study habits.	3.89	1.393	.590	.726
1. When I am studying and there is something around me that may hinder or prevent me from studying, I am able to change this situation (e.g. if there is noise I go to another place, I ask for help from a classmate if I need it, etc.).	4.55	1.367	.424	.769
Learning management				
4. Before starting to study I plan the time needed to achieve the objectives I have set myself.	3.68	1.690	.604	.713
19. When I finish studying or doing an assignment, I check that nothing is missing from what I had planned before I started.	4.25	1.429	.480	.747
7. I establish fixed times for study in my evening or weekend schedule.	3.40	1.663	.553	.728
15. In the exam period, I plan my work in such a way that I have time to study all the exam content.	4.06	1.467	.550	.729
16. In exams, before I start writing I think about how I am going to do it so that I have time to finish it.	3.67	1.536	.331	.777
3. I set goals, I detect what is not working when I am studying and I modify it to improve it.	4.10	1.361	.526	.735

Source: Compiled by author

The KMO test (=.869) and Bartlett's sphericity test ($\chi^2 = 2380.484$; $p < .001$) indicate that the selected sample is adequate to perform the factorisation of the items. An exploratory factor analysis was carried out, extracting three factors through Kaiser's criterion and parallel analysis. The three factors explain 39.90% of the total variance (Table II).

TABLE II. Total variance explained

Initial eigenvalues			
ITEM	Total	% of variance	Accumulated %
1	5.541	26.386	26.386
2	1.491	7.098	33.485
3	1.347	6.416	39.901

Source: Compiled by author

The saturation of the items in each factor and the reliability obtained is described in Table III.

TABLE III. Matrix of rotated components

ITEM	1	2	3
It 6	.669	.028	.082
It 21	.624	-.052	.090
It 18	.573	.266	.325
It 5	.501	.262	.296
It 9	.487	.289	.211
It 20	.445	.111	-.036
It 17	.427	.158	.220
It 1	.352	.241	.261

(Continued)

TABLE III. Matrix of rotated components (Continued)

ITEM	1	2	3
It 12	.120	.669	.066
It 14	.039	.653	.092
It 2	.170	.582	.045
It 13	.252	.553	.312
It 8	-.074	.527	.363
It 10	.346	.441	-.154
It 11	.174	.416	.129
It 7	-.055	-.038	.774
It 4	.286	-.019	.682
It 15	.287	.191	.598
It 19	.252	.283	.582
It 3	.396	.308	.429
It 16	.053	.313	.415

Source: Compiled by author

When analysing the criterial validity, correlations of $r= .558$ ($p<.001$) between the learning to learn competence and the use of a deep focus, $r= -.230$ ($p<.001$) with the use of a shallow focus and $r= .582$ ($p<.001$) and the perception of Self-Efficacy were obtained. The relationship between academic performance and perception of the level of development of the learning to learn competence was $r= .402$ ($p<.001$). These results are in line with those obtained by other authors, as will be seen in the discussion.

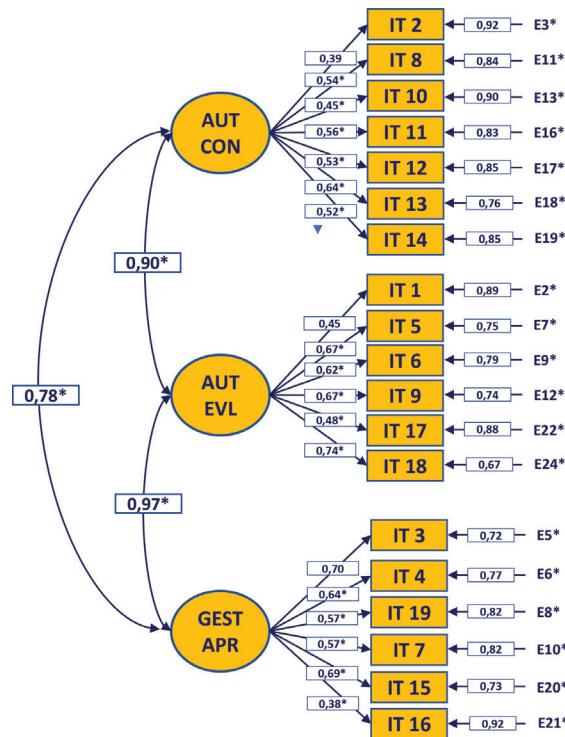
Regarding the confirmatory factor analysis, the goodness-of-fit indices of models A, B and C are presented in Table IV and the diagram of model B is presented in Graph I.

TABLE IV. CFA Model Fit Indices

	Ji²	df	P	Ji²/df	AIC	GFI	CFI	TLI	SRMR	RMSEA
A	445.29	186	<.001	2.39	73.29	.92	.91	.90	.05	.05
B	310.48	149	<.001	2.08	12.88	.94	.94	.93	.04	.04
C	281.14	116	<.001	2.42	49.14	.94	.94	.93	.04	.05

Source: Compiled by author

GRAPH I. Diagram of model B of three related factors



The RMSEA value ($= .04$) was less than $.05$ indicating that the model had a good level of fit (Goh & Yusuf, 2017; Hu & Bentler, 1999). The GFI ($= .94$) and CFI ($= .94$) indices were around $.95$, and can be considered good indicators as they are close to $.95$ (Fan et al., 2016). The TLI index also stood at $.93$. Although it would not meet the $> .95$ criterion (Hu & Bentler, 1999; Xia & Yang, 2019), it would meet that of Goh and Yusuf (2017), as it is $> .90$. The SRMR index would be below the $< .08$ criterion (Cho et al., 2020; Hu & Bentler, 1999), which would make it a good indicator of model fit.

Lastly, the Akaike index score (AIC) was noteworthy. Following the criteria of Burnham and Anderson (1998), the best model is that with the lowest AIC. Thus, in model A, with 21 items, the AIC index was 73.29, in model B with 19 items the AIC was 12.88 and in model C, with 17 items, there was an AIC of 49.14. This was the reason for selecting model B.

The indices calculated in model B of three related factors and 19 items offer a good fit between the theoretical model postulated and it can be considered an acceptable model, there being a satisfactory degree of congruence between the hypothesised theoretical model and what was shown by the empirical data of the sample of secondary school students, which were similar to those obtained in the university sample (Muñoz-San Roque et al., 2016).

Once the model had been defined and in order to respond to the second research objective, the interest of the paper focused on verifying the evolution of the level of development of the competence according to the grade in which the student was enrolled in order to analyse its relationship with age. The results according to the grade indicated that the level of self-perception of the development of the learning to learn competence obtained a higher mean in 1ESO (mean= 4.48), followed by 2ESO (mean= 4.26), 3ESO (mean= 4.12), 4ESO (mean= 4.13) and, finally, baccalaureate (mean= 4.11). The difference was statistically significant ($p < .05$) between 1ESO and the rest of the years except for 2ESO, although the magnitude of the difference was low ($\eta^2 = .02$) (Cohen, 1992). This datum indicates that the perception of the development of this competence decreases as the school years progress, as can be seen in Table V.

TABLE V. Main variables of the research as a function of the year

Variables	Year	Mean	σ	F	η^2	Peer-to-peer comparison
CAaA	1ESO	4.48	0.74	6.57***	.02	1ESO > 3ESO 1ESO > 4ESO 1ESO > BACC
	2ESO	4.26	0.74			
	3ESO	4.12	0.74			
	4ESO	4.13	0.80			
	BACC	4.11	0.79			
	Total	4.20	0.78			
AUTEVL	1ESO	4.51	0.82	7.51***	.03	1ESO > 3ESO 1ESO > 4ESO 1ESO > BACC
	2ESO	4.26	0.88			
	3ESO	4.15	0.87			
	4ESO	4.12	0.92			
	BACC	4.02	0.91			
	Total	4.20	0.90			

(Continued)

TABLE V. Main variables of the research as a function of the year (Continued)

Variables	Year	Mean	σ	F	η^2	Peer-to-peer comparison
AUTCON	1ESO	4.76	0.74	4.53**	.02	1ESO > 2ESO 1ESO > 3ESO 1ESO < 4ESO
	2ESO	4.48	0.80			
	3ESO	4.44	0.73			
	4ESO	4.48	0.81			
	BACC	4.54	0.76			
	Total	4.52	0.78			
GESTAPR	1ESO	4.20	0.97	6.29***	.02	1ESO > 3ESO 1ESO > 4ESO 1ESO > BACC
	2ESO	3.94	0.91			
	3ESO	3.75	1.02			
	4ESO	3.78	1.04			
	BACC	3.75	1.10			
	Total	3.86	1.02			

*p<.05; ** p<.01; ***p<.001

Source: Compiled by author

A second contrast analysis of the means between the first year of ESO (1ESO) and the first year of baccalaureate (Table VI) serves to evaluate the magnitude of the effect of the difference between the first and last

TABLE VI. Learning to learn and its components depending on whether the student is enrolled in 1ESO or in baccalaureate.

Variables	Year	Mean	σ	t	d
CAaA	1ESO	4.49	0.77	3.82***	0.48
	BACC	4.11	0.79		
AUTEVL	1ESO	4.53	0.86	4.67***	0.57
	BACC	4.02	0.91		
AUTCON	1ESO	4.77	0.75	2.42*	0.30
	BACC	4.54	0.76		
GESTAPR	1ESO	4.18	0.98	3.30***	0.41
	BACC	3.75	1.10		

*p<.05; ** p<.01; ***p<.001

Source: Compiled by author

year of the sample. The differences are statistically significant ($p < .05$) and of moderate magnitude in the learning to learn competence (CAaA) ($t = 3.82$; $df = 244$; $p < .001$; $d = 0.48$); in self-assessment of the process (AUTEVL) ($t = 4.67$; $df = 260$; $p < .001$; $d = 0.57$); in self-awareness as a learner (AUTCON) ($t = 2.42$; $df = 255$; $p < .05$; $d = 0.30$) and in learning management (GEST APR) ($t = 3.30$; $df = 261$; $p < .001$; $d = 0.41$).

When the assumptions of homogeneity of variances and normality were not met, the statistics were verified through non-parametric tests (Kruskal-Wallis and Mann-Whitney U), confirming the results presented.

These results show that the students' perception of their level of development in this competence decreases as the academic years progress from 1ESO to the first year of baccalaureate, and that this decrease is considered statistically significant and of moderate magnitude. This leads us to examine this apparent paradox in greater depth, since theory indicates that students in higher school years are able to implement higher order metacognitive and self-regulatory skills (Zimmerman, 2013).

Conclusions

The first objective of this paper was the adaptation of a scale to measure the learning to learn competence in secondary school students which, due to the reliability indices, factorial structure and goodness-of-fit indices presented, can be considered valid and reliable. It is important to note that the factors extracted are in line with the European framework for the assessment of this competence (Hoskins & Fredriksson, 2008). Self-assessment of the process, as a basic metacognitive strategy of the learning to learn competence, appears in the work of Hautamäki et al. (2002) and self-awareness as a learner is identified with what Deakin-Crik et al. (2004) call strategic knowledge in the framework of the elaboration of the European pre-pilot test (Kupiainen et al., 2008). Similarly, the resulting components are in tune with the theoretical framework of the learning to learn competence through self-regulatory competence (Hadwin et al., 2018; Panadero, 2017; Panadero & Tapia, 2014; Salas & Gallardo, 2022; Salmerón & Gutiérrez, 2012; Schunk & Greene, 2018; Torre, 2007; Usher & Schunk, 2018; Winne, 2018; Zimmerman, 2013) and with the European legislative framework (European Commission 2006; European Council, 2018), although with a higher weighting of cognitive factors

with respect to socio-cognitive factors, in accordance with the prevailing legislation at the time of the adaptation of the scale. Further research will include a new dimension encompassing these factors to assess the personal, social and learning to learn competence more broadly.

The relationship observed between learning to learn, self-efficacy and the use of a deep focus has broad correlates with other research (Ardura & Galán, 2019; Biggs, 1987; Cerezo et al., 2019; Kulakow, 2020; Phan, 2011; Ramudo et al., 2017; Schunck & Zimmerman, 1994; Usher & Schunk, 2018), which confirms the validity of the criteria of the scale.

The second aim of the paper, and one of the most significant results of the research, was to demonstrate that, as students progress through their school years and in age, their learning to learn scores decrease significantly. This result was not seen in the European pre-pilot test of learning to learn in Spain (Moreno et al., 2008) or in samples of the university population (Muñoz-San Roque et al., 2016; Torre, 2007). Neither do these differences agree with the theoretical model on the development of self-regulated learning (Zimmerman, 2013), which postulates evolutionary progress from processes of imitation and modelling to others in which the learner can already regulate his or her own learning process. However, several research projects have shown that there is a decrease in learning skills and motivation between lower and higher school years (Dignath & Büttner, 2008; Gaeta, 2013; González Fernández, 2005; Palomo del Blanco, 2014; Pintrich, 2003; Rodríguez Fuentes, 2009; Rosario et al., 2012; Zimmerman & Martínez-Pons, 1990).

This decrease in motivation to learn must be corroborated through research focused on the environmental factors affecting the competence, such as the international study by Stringher (2021), which proposes as a plausible hypothesis that educational systems fail to interest students by using rather traditional methodologies. It is also essential to carry out studies that incorporate the vision of teachers on this aspect, which we consider essential.

It is fundamental that in the development of learning skills there be a parallel between cognitive and metacognitive skills and emotional and motivational skills.

One line of research that would help to explain this decrease in the self-perception of competence as students gradually move up to higher years of secondary education would be to look more closely, through a longitudinal study, at how and when the components of the learning to

learn concept develop evolutionarily, and information should be gathered through performance tests that will give a more objective view of the level of development of the competence.

One limitation of this study is that the instrument designed focused on cognitive and metacognitive aspects. Further research will include the social dimension in the design of instruments for the assessment of this competence, an aspect that is considered fundamental in the new European conceptualisation framework (European Council, 2018) and that has inspired the change in legislation in Spain (LOMLOE, 2020). Another limitation is related to the sample – access to secondary school students is a complicated process, and therefore it was not possible to have a larger and more heterogeneous sample as regards type of school. This fact has meant that the sample used for the adaptation of the instrument and to analyse whether there was a decline in the competence in each school year was the same. However, even with this being the case, it was possible to access information from 1033 students.

Bibliographical References

Abad, F., Olea, J., Ponsoda, V., García, C. (2011). *Medición en ciencias sociales y de la salud Abad*. Síntesis.

Ardura, D., & Galán, A. (2019). The Interplay of Learning Approaches and Self-Efficacy in Secondary School Students' Academic Achievement in Science. *International Journal of Science Education*, 41, 1723-1743. <https://doi.org/10.1080/09500693.2019.1638981>

Bentler, P. M. (1995). *EQS. Structural Equations Program Manual*. Multivariate Software.

Biggs, J. B. (1987). *Student approaches to learning and studying*. Australian Council for Educational Research.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press.

Burnham, K. P., & D. R. Anderson (1998). *Model selection and inference: A practical information-theoretic approach*. Springer-Verlag. <https://doi.org/10.1007/978-1-4757-2917-7>

Byrne, B. M. (2006). *Structural equation modelling with EQS*. Basic Concepts, Applications and Programming. Lawrence Erlbaum Associates.

Caena, F. & Punie, Y. (2019). *Developing a European Framework for the Personal, Social & Learning to Learn Key Competence (LifEComp)*. Publications Office of the European Union. <https://doi.org/10.2760/172528>

Caena, F. & Stringher, C. (2020). Hacia una nueva conceptualización del Aprender a Aprender. *Aula Abierta*, 49(3), 199-216. <https://doi.org/10.17811/rifie.49.3.2020.225-244>

Cerez, R., Fernández, E., Amieiro, N., Valle, A., Rosario, P., & Núñez, J. C. (2019). El papel mediador de la autoeficacia y la utilidad entre el conocimiento y el uso de estrategias de autorregulación del aprendizaje. *Revista de Psicodidáctica*, 24(1), 1-8. <https://doi.org/10.1016/j.psicod.2018.08.001>

Cho, G., Hwang, H., Sarstedt, M., & Ringle, C. M. (2020). Cutoff criteria for overall model fit indexes in generalized structured component analysis. *Journal of Marketing Analytics*, 8, 189–202. <https://doi.org/10.1057/s41270-020-00089-1>

European Commission (2006). Recomendación del Parlamento Europeo y el Consejo sobre las Competencias Clave para el Aprendizaje Permanente (2006/962/EC). *Diario Oficial de la Unión Europea*, Serie L, 394, pp. 10-18. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:ES:PDF>

European Council (2000). *Conclusiones de la Presidencia del Consejo Europeo celebrado en Lisboa*, los días 23 y 24 de marzo de 2000. https://www.europarl.europa.eu/summits/lis1_es.htm#

European Council (2018). Recomendación del Consejo de 22 de mayo de 2018 relativa a las competencias clave para el aprendizaje permanente. (2018/C 189/01). Diario oficial de la Unión Europea, Serie C, nº 189/1, de 4 de junio de 2018. [https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018H0604\(01\)&from=SV](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32018H0604(01)&from=SV)

Deakin, R., Broadfoot, P. & Claxton, G. (2004). Developing an effective lifelong learning inventory: The ELLI project. *Assessment in Education*, 11(3), 247-272. <http://dx.doi.org/10.1080/0969594042000304582>.

Deakin Crick, R. (2014). Learning to learn: a complex systems perspective. In R. Deakin Crick, C. Stringher & K. Ren (Eds.) *Learning to Learn. International Perspectives from Theory and Practice* (pp. 69–90). Routledge.

De Carvalho, J. and Chima, F.O. (2014). Applications of structural equation modeling in social sciences research, *American International Journal of Contemporary Research*, 4(1), 6-11. http://www.aijcrnet.com/journals/Vol_4_No_1_January_2014/2.pdf

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta - analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231-264. <http://10.1007/s11409-008-9029-x>

Efkrides, A. (2011). Interactions of metacognition with emotion and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6-25. <https://doi.org/10.1080/00461520.2011.538645>

Elshout-Mohr, M., Meijer, J., Oostdam, R. & Van Gelderen, A. (2004). *CCST: A test for cross-curricular skills*. SCO – Kohnstamm Institution, University of Amsterdam.

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, 5(19), 1-12. <https://doi.org/10.1186/s13717-016-0063-3>

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33. <https://www.redalyc.org/articulo.oa?id=77812441003>

Fredriksson, U. Hoskins, B. (2006a). *Learning to learn network meeting Report*. CRELL/JRC.

Fredriksson, U. Hoskins, B. (2006b). *Learning to learn network meeting Report from the second meeting of the network*. CRELL/JRC.

Fredriksson, U. Hoskins, B. (2007). *Learning to learn network meeting Report from the third meeting of the network*. CRELL/JRC.

Gaeta, M.L. (2013). Promoción del aprendizaje autorregulado en la enseñanza secundaria: un estudio comparativo. *Revista Qurriculum*, 26, 161-176. <http://riull.ull.es/xmlui/handle/915/4444>

Garcia-Garcia, F. J., López-Francés, I., Gargallo-López, B., & Pérez-Pérez, C. (2022). Validación de contenido de la competencia “aprender a aprender” en los grados universitarios. *Revista de Investigación Educativa*, 40(2), 513-530. <https://doi.org/10.6018/rie.466271>

Goh, P. S., & Yusuf, Q. (2017). Validation of the Malaysian Version of the Teacher Education Program Coherence Questionnaire. *Australian Journal of Teacher Education*, 42(12). <https://doi.org/10.14221/ajte.2017v42n12.3>

González Fernández, A. (2005). *Motivación académica. Teoría, aplicación y evaluación*. Pirámide.

González Geraldo, J. L., Del Rincón Igea, B., & Bayot Mestre, A. (2010). Enfoques de aprendizaje y rendimiento académico en educación

secundaria. *Revista Galego-Portuguesa de Psicoloxía e Educación*, 18(1), 211-226. <https://core.ac.uk/download/pdf/61903119.pdf>

González-Montesinos, M. & Backhoff, E. (2010). Validación de un cuestionario de contexto para evaluar sistemas educativos con Modelos de Ecuaciones Estructurales. *RELIEVE*, 16(2), 1-17. <https://doi.org/10.7203/relieve.16.2.4133>

Hadwin, A. F., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed.) (pp. 80-102). Routledge.

Hair, J., Anderson, R., Tatham, R. & Black, W. (1998). *Multivariate data analysis*. 5th Edition, Prentice Hall, New Jersey.

Hautamäki, J., Arinen, P., Eronen, S., Hautamäki, A., Kupianien, S., Lindblom, B., Niemivirta, M., Pakaslahti, L., Rantanen, P. & Scheinin, P. (2002). *Assessing Learning-to-Learn: A Framework*. Centre for Educational Assessment, Helsinki University / National Board of Education.

Hernández & Franco V. Gonzalo Misol, I. (2009). Validación de un sistema de indicadores para valorar la integración socioeducativa de alumnos emigrantes. En A. Boza Carreño, J. M. Méndez Garrido, M. Monescillo Palomo, M. Toscano Cruz, M. Aguaded Gómez, J. A. Ávila Fernández, J. Tello Díaz, M. Salas Tenorio. (Coords.). *Educación, investigación y desarrollo social. Actas del XIV Congreso Nacional de modelos de Investigación Educativa (AIDIPE)* (pp. 239-280). Universidad de Huelva.

Hoskins, B. & Fredriksson, U. (2008). *Learning to learn: What is it and can it be measured?* Office for Official Publications of the European Communities. <https://doi.org/10.2788/83908>

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. <https://doi.org/10.1080/10705519909540118>

Hutmacher, W. (1997). Key competencies in Europe. *European Journal of Education*, 32 (1), 45-58. <https://doi.org/10.1111/1467-3435.00004>.

Jöreskog, K. G. (1970). A general method for estimating a linear structural equation system. *ETS Research Bulletin Series*, i-41. <https://doi.org/10.1002/j.2333-8504.1970.tb00783.x>

Kember, D., Biggs, J. & Leung, D. (2004). Examining the multidimensionality of approaches to learning through the development of

a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, 74, 261-280. <https://doi.org/10.1348/000709904773839879>

Kline, R.B. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Kulakow, S. (2020). How autonomy support mediates the relationship between self-efficacy and approaches to learning. *The Journal of Educational Research*, 113(1), 13-25. <https://doi.org/10.1080/00220671.2019.1709402>

Kupianen, S., Hautamäki, J., & Rantanen, P. (2008). *EU pre-pilot on learning to learn: Report on the compiled data*, 2008-1190/001-001 TRA-TRINDC. European Commission.

LOE (2006). Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boletín Oficial del Estado (España)*, 4 de mayo de 2006, 106, 17158-17207. <http://www.boe.es/boe/dias/2006/05/04/pdfs/A17158-17207.pdf>

LOMCE (2013). Ley Orgánica Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. *Boletín Oficial del Estado (España)*, 10 de diciembre de 2013, 295, 1-63. <https://www.boe.es/buscar/pdf/2013/BOE-A-2013-12886-consolidado.pdf>

LOMLOE (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. Boletín Oficial del Estado (España), 30 de diciembre de 2020, 34, 122868-122953. <https://www.boe.es/eli/es/lo/2020/12/29/3>.

Martín-Alonso, J. F. (2021). La competencia de aprender a aprender. Diseño y validación de un instrumento para medir la percepción de su desarrollo en alumnos de secundaria [Tesis doctoral, Universidad Pontificia Comillas]. Repositorio institucional de la Universidad Pontificia Comillas. <http://hdl.handle.net/11531/62491>

Martín Alonso, J. F., & Muñoz San Roque, I. (2022). LifeComp en España: Retrospectiva. *Cuadernos de pedagogía*, (528), 15.

Martín, E. & Moreno, A. (2007). *Competencia para aprender a aprender*. Alianza Editorial.

Martínez-Abad, F., & Rodríguez-Conde, M. J. (2017). Comportamiento de las correlaciones producto-momento y tetracórica-policórica en escalas ordinales: un estudio de simulación. *RELIEVE - Revista Electrónica De Investigación Y Evaluación Educativa*, 23(2). <https://doi.org/10.7203/relieve.23.2.9476>

Martínez-Fernández, R. (2007). Concepción de aprendizaje y estrategias metacognitivas en estudiantes universitarios de psicología. *Anales de Psicología*, 3(1), 7-16. <https://revistas.um.es/analesps/article/view/23261>

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: Lawrence Erlbaum Associates, Inc.

Moreno, A. (2002). La evaluación de las habilidades metacognitivas. En A. Marchesi y E. Martín (Eds.), *Evaluación de la Educación Secundaria*. pp. 119-136). SM.

Moreno, A., Cercadillo, L. & Martínez, M. (2008). *Learn European Project. Pre-pilot study national report*. Instituto de Evaluación (MEC).

Moreno, A. & Martín, E. (2007). The development of learning to learn in Spain. *The curriculum Journal*, 18(2):175-193. <https://doi.org/10.1080/09585170701446028>

Muñoz-San Roque, I., Martín-Alonso, J. F., Prieto-Navarro, L., & Urosa-Sanz, B. (2016). Autopercepción del nivel de desarrollo de la competencia de aprender a aprender en el contexto universitario: propuesta de un instrumento de evaluación. *Revista de Investigación Educativa*, 34(2), 369-383. <https://doi.org/10.6018/rie.34.2.235881>

Palomo del Blanco, M. (2014). El autoconcepto y la motivación escolar: una revisión bibliográfica. *International Journal of Developmental and Educational Psychology*, 6(1), 221-228. <https://doi.org/10.17060/ijodaep.2014.n1.v6.737>

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>

Panadero, E., & Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. *Anales de Psicología*. 3. 450-462. <http://dx.doi.org/10.6018/analesps.30.2.167221>

Phan, H. P. (2011). Interrelations between self-efficacy and learning approaches: a developmental approach. *Educational Psychology*, 31(2), 225-246. <https://doi.org/10.1080/01443410.2010.545050>

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686. <https://doi.org/10.1037/0022-0663.95.4.667>

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 452-502). Academic Press.

Ramudo, I., Brenlla, J. C., Barca, A., & Peralbo, M. (2017). Enfoques de aprendizaje, autoeficacia y rendimiento académico en el alumnado de bachillerato: Implicaciones para la enseñanza. *Revista de Estudios e Investigación en Psicología y Educación. Extr.*(1), A1-139. <https://doi.org/10.17979/reipe.2017.0.01.2435>

Rodríguez Fuentes, G. (2009). Motivación, estrategias de aprendizaje y rendimiento académico en estudiantes de E.S.O. [Tesis doctoral, Universidad de la Coruña]. RUC. <https://ruc.udc.es/dspace/handle/2183/10117>

Rojas-Torres, L. (2020). Robustez de los índices de ajuste del análisis factorial confirmatorio a los valores extremos. *Revista de matemática: teoría y aplicaciones*, 27(2), 383–404. <https://doi.org/10.15517/rmta.v27i2.33677>

Rosário, P., Lourenço, A., Paiva, M. O., Núñez, J. C., Pienda, J. A. G., & Valle, A. (2012). Autoeficacia y utilidad percibida como condiciones necesarias para un aprendizaje académico autorregulado. *Anales de Psicología/Annals of Psychology*, 28(1), 37-44. <https://www.redalyc.org/pdf/167/16723161005.pdf>

Sala, A., Punie, Y., Garkov, V., & Cabrera Giraldez, M. (2020). *LifeComp: The European Framework for Personal, Social and Learning to Learn Key Competence*. Publications Office of the European Union

Salas Guadiana, E. A., & Gallardo Córdova, K. E. (2022). Evolución de la conceptualización de la autorregulación en Orientación Educativa: una revisión sistemática de literatura. *REOP - Revista Española De Orientación Y Psicopedagogía*, 33(2), 23–44. <https://doi.org/10.5944/reop.vol.33.num.2.2022.34358>

Salmerón Pérez, H. & Gutierrez Braojos, C. (2012). La competencia de aprender a aprender y el aprendizaje autorregulado. Posicionamientos teóricos. *Revista de Currículum y Formación de Profesorado*, 16(1), 1-9. <https://dialnet.unirioja.es/servlet/articulo?codigo=4643314&orden=1&info=link>

Salmerón-Pérez, H., Gutierrez-Braojos, C., Fernández-Cano, A., & Salmerón-Vilchez, P. (2010). Aprendizaje Autorregulado, Creencias De Autoeficacia Y Desempeño en La Segunda Infancia. *RELIEVE - Revista*

Electrónica de Investigación y Evaluación Educativa, 16(2), 1-18. <https://doi.org/10.7203/relieve.16.2.4136>

Schunk, D. H. & Greene, J. A. (Eds.). (2018). *Handbook of self-regulation of learning and performance*. Routledge, Taylor & Francis Group.

Schunk, D.H., & Zimmerman, B. (1994). *Self-regulation of learning and performance: issues and educational applications*. Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, 32 (4), 195–208. https://doi.org/10.1207/s15326985ep3204_1

Stringher, C. (2014). What is learning to learn? A learning to learn process and output model. En R. Deakin Crick, C. Stringher, & K. Ren (Eds.), *Learning to learn* (9-32). Routledge.

Stringher, C. (2021). Cos'è l'Apprendere ad Apprendere? Alcune risposte a livello istituzionale e teorico. In C. Stringher (Ed.), *Apprendere ad apprendere in prospettiva socioculturale: rappresentazioni dei docenti in sei Paesi* (pp. 19–49). Franco Angeli.

Stringher, C., Brito Rivera, H. A., Patera, S., Silva Silva, I., Castro Zubizarreta, A., Davis Leme, C., Torti, D., Huerta, M. C. & Scrocca, F. (2021) Learning to learn and assessment: Complementary concepts or different worlds? *Educational Research*, 63(1), 26-42. <http://10.1080/00131881.2021.1871576>

Suárez, J. M., Anaya, D. & Fernández, A. P. (2005). Un modelo sobre la determinación motivacional del aprendizaje autorregulado. *Revista de Educación*, 338, 295-308. <https://doi.org/10.4438/1988-592X-RE-2005-338-016>

Torre Puente, J.C. (2007). *Una triple alianza para un aprendizaje universitario de calidad*. Universidad Pontificia Comillas.

Usher, E. L., & Schunk, D. H. (2018). Social cognitive theoretical perspective of self-regulation. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed.) (p. 19–35). Routledge.

Valle, J.M. (2020). ¿Nuevas competencias para la vida o competencias para una nueva vida? La renovada visión de la Unión Europea en las Competencias Clave 2.0 (2018). En J. Moya e J. Valle (Coords.), I., *La reforma del currículo escolar: ideas y propuestas* (pp. 83-110). ANELE.

Villardón-Gallego, L., Yániz, C. Achurra, C. Iraurgi, J., & Aguilar, C. (2013). Learning competence in university: development and structural

validation of a scale to measure/ La competencia para aprender en la universidad: desarrollo y validación de un instrumento de medida. *Revista de Psicodidáctica*, 18(2), 357-374. <http://doi.org/10.1387/RevPsicodidact.6470>

Whitebread, D., Bingham, S., Grau, V., Pino-Pasternak, D. & Sangster, C. (2007). The development of metacognition and self-regulated learning in young children: The role of collaborative and peer-assisted learning, *Journal of Cognitive Education and Psychology*, 6(3), 433–455. <https://doi.org/10.1891/194589507787382043>

Winne, P. H. (2018). Cognition and metacognition within self-regulated learning In. D.H., Schunk & J. A., Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd Ed.). (pp. 52-64). Routledge.

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 Methods*, 51(1), 409–428. <https://doi.org/10.3758/s13428-018-1055-2>

Zimmerman, B. J. (2013). From Cognitive Modeling to Self-Regulation: A Social Cognitive Career Path. *Educational Psychologist*, 48(3), 135-147. <https://doi.org/10.1080/00461520.2013.794676>

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective, *Educational Psychologist*, 30(4), 217-221. https://doi.org/10.1207/s15326985ep3004_8

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51. <https://doi.org/10.1037/0022-0663.82.1.51>.

Contact address: Isabel Muñoz San Roque. Departamento de Educación, Métodos de Investigación y Evaluación de la Universidad Pontificia Comillas de Madrid. Universidad Comillas, 3 28049 Madrid. E-mail: isabelmsanroque@comillas.edu

Challenges in teaching practices for the incorporation of sustainable development approach in Compulsory Secondary Education

Retos en las prácticas docentes para la incorporación del enfoque del desarrollo sostenible en la Educación Secundaria Obligatoria

<https://doi.org/10.4438/1988-592X-RE-2023-401-583>

Belén Sáenz-Rico de Santiago

<https://orcid.org/0000-0001-9207-889X>

Universidad Complutense de Madrid

M^a del Rosario Mendoza Carretero

<https://orcid.org/0000-0002-6924-3196>

Universidad Complutense de Madrid

Raúl García Medina

<https://orcid.org/0000-0003-4367-3160>

Universidad Complutense de Madrid

Mercedes Sánchez Sáinz

<https://orcid.org/0000-0002-7547-4724>

Universidad Complutense de Madrid

Abstract

The purpose of this study¹ is to investigate the relationship between the degree of knowledge of the teachers of the Compulsory Secondary Education stage as regards sustainability and their teaching practice within the framework

¹ This research is framed within the: Eco-Social Literacy Project: a central element in the processes of curricular sustainability for the achievement of the 2030 Agenda (SDGs), in initial teacher training (PRAD-ODS) RTI2018-095746-B-I00, funded by the Ministry of Science, Innovation and Universities.

of the 2030 Agenda, as well as to identify the challenges they face to promote the sustainability approach in the classroom. **Methodology:** the data have been collected through a questionnaire sent to schools of Compulsory Secondary Education, validated by experts and with a reliability of $\alpha = .89$. It has a non-probability sampling for convenience ($N = 826$), obtained using the snowball technique. A descriptive and inferential analysis was carried out, using an analysis of non-parametric tests. **Results:** The teachers who teach in centres that carry out projects and activities to raise awareness and awareness in the field of sustainability, have a greater degree of knowledge of both the 2030 Agenda and the SDGs and incorporate it into their subjects and into the development of competencies. ; The main challenges referenced by the teaching staff respond to the absence of the sustainable development approach in the organizational documents, time management for curricular innovation, and the need for specific training. **Conclusion:** The relationship between sensitization and awareness of teachers and their teaching practice oriented towards sustainability is confirmed. It is necessary to reinforce the training of teachers to increase their knowledge of the great challenges of the planet and the role of education in the transformation towards more sustainable societies.

Keywords: compulsory secondary education, eco-social literacy, teacher education, teaching practice, sustainable development.

Resumen

Este estudio² tiene como propósito indagar sobre la relación entre el grado de conocimiento del profesorado de la etapa de Educación Secundaria Obligatoria en materia de sostenibilidad y su práctica docente en el marco de la Agenda 2030, así como identificar los desafíos a los que se enfrenta para promover el enfoque de la sostenibilidad en el aula. **Metodología:** los datos se han recogido mediante un cuestionario enviado a centros educativos de Educación Secundaria Obligatoria, validado por personas expertas y con una fiabilidad de $\alpha = .89$. Cuenta con un muestreo no probabilístico por conveniencia ($N = 826$), obtenido mediante técnica de bola de nieve. Se realizó un análisis descriptivo e inferencial, utilizando un análisis de pruebas no paramétricas. **Resultados:** El profesorado que ejerce su docencia en centros que desarrollan proyectos y actividades de sensibilización y concienciación en materia de sostenibilidad, tiene mayor grado de conocimiento tanto de la Agenda 2030 como de los ODS y lo incardina en sus materias y en el desarrollo de competencias; los principales desafíos referenciados por el profesorado responden a la ausencia del enfoque de desarrollo sostenible en los

² Esta investigación está enmarcada en el proyecto: La alfabetización ecosocial: un elemento central en los procesos de sostenibilización curricular para el logro de la Agenda 2030 (ODS), en la formación inicial del profesorado (PRAD-ODS) RTI2018-095746-B-I00, financiado por el Ministerio de Ciencia, Innovación y Universidades.

documentos organizativos, a la gestión del tiempo para la innovación curricular, y a la necesidad de formación específica. Conclusión: Se confirma la relación entre sensibilización y concienciación del profesorado y su práctica docente orientada a la sostenibilidad. Es preciso reforzar la formación del profesorado para incrementar su conocimiento ante los grandes retos del planeta y el papel de la educación en la transformación hacia sociedades más sostenibles.

Palabras clave: educación secundaria obligatoria, alfabetización ecosocial, formación de docentes, práctica pedagógica, desarrollo sostenible.

Introduction

The Earth Summit, held in Rio de Janeiro in 1992, recognised that education had an essential role to play in tackling various issues, such as socio-ecological issues (Murga-Menoyo, 2021), citizenship issues, and progress towards sustainability (Calero et al., 2019; Vilches & Gil, 2012).

The Decade of Education for Sustainable Development (2005-2014), proclaimed by the United Nations (UN) (Gil et al., 2006; United Nations Educational, Scientific and Cultural Organization [UNESCO], n.d.), aimed to engage the population in the transition towards sustainability (Calero et al., 2019), using education.

Education for sustainability uses various strategies to raise awareness of environmental and social justice issues (Fernández & Gutiérrez, 2014). It aims to solve “social, economic and ecological problems” (Fundación Benéfico-Social Hogar del empleado [FUHEM], 2018, p.9), and is involved in transforming people, improving their self-perception of eco-dependence and “the deep interdependencies that allow us to be alive” (FUHEM, 2018, p.9). It seeks to “develop competences for participation in improvement actions within the community” (Fernández & Gutiérrez, 2014, p.28).

Education is an engine for change in society (Fernández & Casado, 2017; Reyes, 2010), and teachers, regardless of their discipline or academic level (Calero et al., 2019), have the role of facilitating the acquisition of an adequate vision of the problems and challenges affecting humanity. With this support, citizens will be able to make evidence-based decisions (Alcalá et al., 2020) and “take responsible actions in relation to environmental integrity, economic viability and a just society, for current and future generations, while respecting cultural diversity” (Murga-Menoyo, 2021, p. 114).

It is therefore necessary to disseminate the model of education for sustainable development by addressing effective curricular sustainability processes. This is achieved when the principles and values of the approach permeate all teaching spheres and are incorporated into the curricula of the different educational levels (Azcárate et al., 2016), as well as in teacher training (Álvarez-García et al., 2018; Calero et al., 2019), which Valderrama et al. (2020) describe as scarce and inconsistent.

In fact, target 4.7, one of the targets of SDG 4, quality education, of the 2030 Agenda approved by the United Nations (UN, 2015), defends the need for students to acquire the theoretical and practical knowledge necessary to promote sustainable development (UN, 2015). Similarly, Royal Decree 217/2022, of 29 March, which establishes the organisation and minimum teaching requirements for Compulsory Secondary Education, advocates promoting sustainable development through safe actions for students, “responding to the need to promote competent, autonomous, meaningful and reflective learning in all subjects” (BOE, p.4). Undoubtedly, teacher training must contribute to achieving this goal (Calero et al., 2019).

This leads us to ask and address the following questions: Is teacher training in sustainability adequate for teachers to take up this challenge? What knowledge and skills might teachers need to develop awareness-raising actions in favour of education for sustainable development in compulsory secondary education?

Teacher training for making the Compulsory Secondary Education curriculum more sustainable

Curricular sustainability is a cross-cutting pedagogical process that enables students to acquire the necessary competences and make a positive contribution to sustainable development. This requires that “the principles, values and procedures of education for sustainable development” (Murga-Menoyo & Novo, 2014, p.167) be incorporated into teacher training; only then will they become a reality in the classroom.

For the sake of coherence, making the curriculum sustainable requires, among other aspects, the involvement of teachers to generate changes that contribute to overcoming the challenges found and try to favour the immersion of students in the culture of sustainability (Alcalá et al., 2020;

Valderrama et al., 2020). To this end, teachers must develop and strengthen the sustainability competences (Brundiers et al., 2021; UNESCO, 2017) necessary for students, so that citizens feel concerned about, integrate and respond to the current eco-social challenges arising from the situation of the planet. Therefore, it has to raise questions and promote strategies that lead to the search for “critical and creative responses to socio-environmental conflicts, generating a social culture based on cooperation and communication, responsibility and participation” (Fernández & Gutiérrez, 2014, p.150).

This demand is not new; studies such as that of Vilches and Gil (2013) already pointed out the need to include the sustainability approach in the curriculum and in teacher training. Even years earlier, in Order ECI/3858/2007, of 27 December, which established the requirements for the verification of official university degrees that enabled the exercise of the professions of Compulsory Secondary Education and Baccalaureate, Vocational Training and Language Teaching (BOE, 2007), the construction of a sustainable future was already mentioned as one of the aims of learning spaces:

Designing and developing learning spaces with special attention to equity, emotional and values education, equal rights and opportunities between men and women, citizenship training and respect for human rights that facilitate life in society, decision-making and the construction of a sustainable future (p.3).

Even so, at present, teachers continue to face challenges in their teaching professionalisation, which are related to sustainability issues and linked to their training, and which can make it difficult for them to advance in the process of curricular sustainability (Alcalá et al., 2020). They seem to feel that they do not have enough knowledge or tools to be able to contribute to a transformative education that favours a change in citizens' attitudes and behaviour in coherence with the sustainability approach (Alcalá et al., 2020).

The training of future teachers, despite the scientific background it provides (Murga-Menoyo & Novo, 2014), does not seem adequate for the transition towards sustainable societies. This could be due to several factors. On the one hand, there are gaps in the contents that favour its acquisition (Aznar et al., 2017; Vilches and Gil, 2012), although they do incorporate the training of sustainability competences within subjects and

disciplines. This hypothesis coincides with the results of other research, such as that of Filho et al. (2017), which warns of the teachers' lack of knowledge about how to environmentalise the curriculum of their subject. On the other hand, it could also be due to the teachers' own attitude towards their new role.

This role has changed due to the transformation of education. Teachers are no longer seen as mere transmitters of knowledge (Álvarez-García et al., 2018; Bermúdez & Lía, 2008; Southwell, 2013); they now have to respond to new literacies, which are interconnected and more complex. To address them, it is necessary to strengthen its training; to become literate and trained (Álvarez-García et al., 2018) in order to be able to implement a model of education for sustainability (McKeown-Ice, 2000; Olaskoaga-Larrauri et al., 2021).

Vilches and Gil (2013) suggested that teachers should be imbued with the culture of sustainability through their own teaching-learning process, so that they can then extrapolate it to the classroom with their students. This strategy, for example, could encourage education for sustainable development in secondary school classrooms. However, in the words of Aznar et al. (2017), "sustainability (...) is still far from achieving a reorientation of the curriculum in relation to sustainable development" (p. 227).

For all these reasons, it is considered relevant to identify the degree of knowledge in sustainability that teachers have, as agents of training and transformation, given the multiplier effect they can have on students during the teaching-learning process, in order to move towards a more sustainable society. Education for sustainability is a fundamental dimension of teacher training.

Education for sustainable development in the curriculum in compulsory secondary education

Acquiring competence in sustainability is a fundamental objective for sustainable development in education (Scharenberg et al., 2021). In Spain, it is anchored in the curricula, covering all educational stages corresponding to basic education.

Organic Law 3/2020 of 29 December, which amends Organic Law 2/2006 of 3 May on Education, states in its preamble that "education for

sustainable development and global citizenship must be included in the educational plans and programmes of all compulsory education, incorporating knowledge, skills, values and attitudes" (BOE, 2020, p.122871) so that people can make informed decisions, taking an active role in tackling and solving problems that concern citizenship.

In the organisation of the four years of this educational stage, in all subjects, in a cross-cutting manner, it is prescribed that education for sustainable development will be promoted, closely related to the critical and scientific spirit, emotional and values education, education for health –including affective-sexual education–, gender equality, mutual respect and cooperation among equals (BOE, 2020).

In this sense, FUHEM (2018) argues that the eco-social approach could be achieved through the curricular areas related to Biology and Geology, Physics and Chemistry, Geography and History, Economics, Introduction to Entrepreneurship and Business, and Eco-Social Values, because if students acquire basic knowledge and skills that enable them to understand the eco-social situation they are living in, they will be able to recognise the eco-dependence of human beings.

This would help to respond to the need for the formation of a planetary citizenship (Murga-Menoyo & Novo, 2017) and the development of the sustainability competences established by UNESCO (2017), so that citizens are capable of "reflection on one's own actions, bearing in mind their social, cultural, economic and environmental effects" (UNESCO, 2017, p.7) and can thus contribute to achieving the necessary transformation.

Teachers therefore need to be trained in sustainability to acquire a commitment and address what the curriculum establishes in their teaching practice, with an eco-social perspective. As FUHEM (2018) points out, "when educating, we should aim to help the people we train acquire the skills they will need throughout their lives" (p.12).

There are various teaching practices, such as that proposed in the study by Scharenberg et al. (2021), in which students acquire sustainability competences. This research shows that teachers' attitudes and knowledge about education for sustainable development are significant predictors of these outcomes. Therefore, the more knowledge teachers have, the more training, awareness and sensitivity towards sustainability, the more sustainability-related knowledge students will acquire.

Objective

In line with the above, the purpose of the research presented here is to investigate the relationship between the level of sustainability knowledge on the part of teachers of Compulsory Secondary Education (hereinafter ESO) and their teaching practice in the framework of the 2030 Agenda, as well as to identify the challenges faced by teachers in promoting the sustainability approach in the classroom.

Method

This quantitative, cross-sectional, multi-centre study was carried out at national level. Considering the following hypotheses:

- H1: The degree of knowledge that teachers have about the 2030 Agenda and the SDGs influences the implementation of programmes and activities related to environmental, social and economic sustainability in secondary schools.
- H2: Teachers' awareness and awareness of the sustainability approach reduces possible barriers to its inclusion in the teaching-learning process in secondary schools.

Sample

This research is based on non-probabilistic convenience sampling using a snowball technique. The sample consists of a total number of 826 teachers working at the Compulsory Secondary Education stage. 63.9% ($n = 528$) of the cohort are female, 35.1% ($n = 290$) are male and 1% ($n = 8$) are of other gender identities. The teachers in this study are aged 20-30 (6.4%; $n = 53$), 30-40 (19.9%; $n = 164$), 40-50 (35.1%; $n = 290$), 50-60 (34.4%; $n = 284$) and over 60 (4.2%; $n = 35$). They teach in public (91.8%; $n = 758$), private (1%; $n = 8$) and state-subsidised schools (7.3%; $n = 60$). The participation of teaching staff by Autonomous Community (hereinafter, AC) is shown in Table I.

TABLE I. Sample representation by Autonomous Community (AC)

Autonomous Community	Representation
Andalusia	10.3% (n = 85)
Aragon	1.8% (n = 15)
Balearic Islands	1.3% (n = 11)
Canary Islands	2.7% (n = 22)
Cantabria	1% (n = 8)
Castile and León	7.4% (n = 61)
Castile-La Mancha	6.5% (n = 54)
Catalonia	7.1% (n = 59)
Ceuta	0.4% (n = 3)
Community of Madrid	18.5% (n = 153)
Valencian Community	15.5% (n = 128)
Extremadura	1.2% (n = 10)
Galicia	8% (n = 66)
La Rioja	0.6% (n = 5)
Navarre	2.7% (n = 22)
Basque Country	3.8% (n = 31)
Principality of Asturias	8% (n = 66)
Region of Murcia	3.3% (n = 27)

Source: Compiled by author

Variables and Instrument

The variables that this research subjects to analysis are: years of teaching, tenure, autonomous community, degree of knowledge of the 2030 Agenda, degree of knowledge of the SDGs³, curriculum competencies, ESO subjects, the difficulties faced by teachers in implementing the 2030

³ It should be clarified that we differentiate between these two variables as the 17 Sustainable Development Goals are part of the 2030 Agenda for Sustainable Development. This makes the 2030 Agenda the framework for people and the planet in the 21st century, and the SDGs and their targets facilitators of the implementation of the 2030 Agenda, but not the only reference, as the Agenda incorporates, among others, the Universal Declaration of Human Rights and the Addis Ababa Action Agenda (UN, 2015).

Agenda and the SDGs in the classroom, and projects and activities to raise awareness of the SDGs, the latter being configured as the dependent variable.

A 'Questionnaire to inquire about ESO teachers' eco-social literacy' was developed using the Google Forms platform. This questionnaire is made up of a total of 14 items, which could be answered by single, multiple or Likert-type responses with an ordinal measurement level of five points. It was administered in 18 Autonomous Regions (1. Andalusia; 2. Extremadura; 13. Galicia; 14. La Rioja; 15. Navarre; 16. Basque Country; 17. Principality of Asturias; 18. Region of Murcia), sent anonymously to different secondary schools during the 2020-2021 academic year.

The questionnaire was validated by 10 experts from different fields: (2) sustainable development, (2) secondary school teachers, (2) the Conference of Rectors of Spanish Universities (CRUE), (1) university experts, (2) questionnaire experts and (1) civilians. The reliability analysis showed good results for the instrument as a whole (TABLE II):

TABLE II. Reliability analysis of the survey

Ordinal Reliability	Raw_alpha (.893)
Ordinal Omega	Omega_h (.715)
	Omega.lim (.773)
	Alpha (.900)
	.938

Source: Compiled by author

Procedure

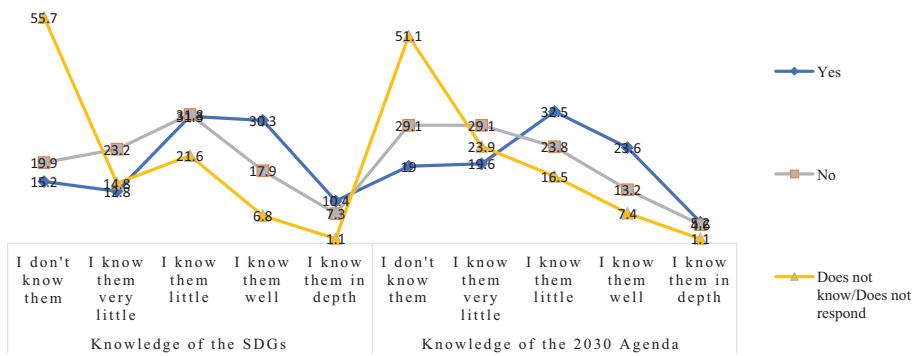
A descriptive and inferential analysis was performed, using a non-parametric test analysis using Pearson's Chi-Square ($p < .05$) through the statistical programme JAMOVI 2.2.5 and IBM Statistical Package for Social Sciences (SPSS) version 27.

Results

It is important to begin by highlighting that, in the questionnaire on eco-social literacy administered, 94.2% of the teachers who responded considered that the basic concepts of environmental, social and economic sustainability should be part of the ESO curriculum.

In response to H1, we can say that the degree of knowledge of the 2030 Agenda ($p = .000$) and the SDGs ($p = .000$) on the part of teachers correlates significantly ($p = < .05$) with the development of awareness-raising projects and activities by the school. Teachers who teach in schools that develop awareness-raising and sensitisation projects and activities have a higher level of knowledge about the 2030 Agenda (17.4%; $n = 144$) and the SDGs (24.5%; $n = 203$) compared to those who do not know whether their school develops projects with this approach (22.7%; $n = 188$). This shows that developing projects linked to the SDGs at the school where they teach leads to greater awareness of moving towards education for sustainable development in the classroom in line with H2 (FIGURE I).

FIGURE I. Development of projects and activities at school and level of awareness of the 2030 Agenda and SDGs



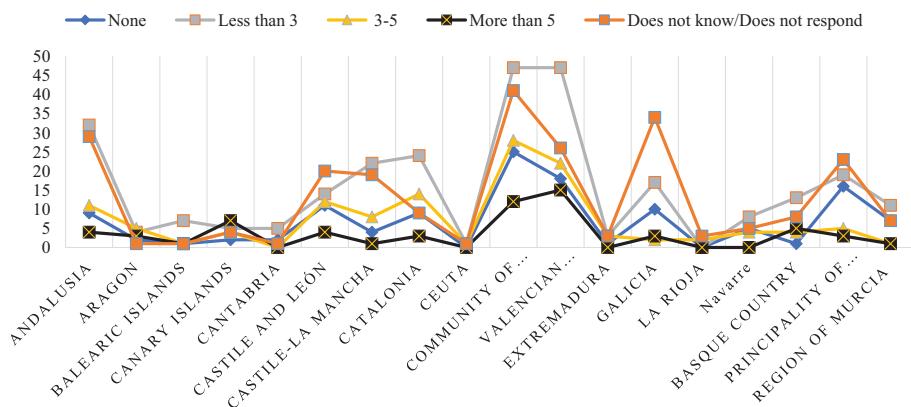
Source: Compiled by author

In turn, it is worth mentioning that there is a significant relationship ($p = <.05$) between the AC and the degree of knowledge that teachers have of the 2030 Agenda ($p = .001$) and the SDGs ($p = .000$), as well as between the AC and the development of projects and activities to raise

awareness of the importance of the SDGs ($p = .000$). 60.4% ($n = 499$) indicate that awareness-raising and sensitisation projects and activities on the importance of the SDGs are developed in their school, compared to 18.3% ($n = 151$) who say no, and 21.3% ($n = 176$) who do not know or do not answer.

Regarding the number of projects and activities implemented in the centre, their distribution is as follows: 14.9% ($n = 123$) indicate that none are developed, 33.8% ($n = 279$) indicate between one and two, 15.4% ($n = 127$) between three and five, 7.7% ($n = 64$) more than six, and 28.2% ($n = 233$) do not know or do not answer. In this sense, according to what was reported by the participants, the centres that develop the most projects and activities are those located in Andalusia, Castile and Leon, Castile-La Mancha, Catalonia, Community of Madrid, Community of Valencia, Basque Country and Principality of Asturias (FIGURE II).

FIGURE II. No. of projects and activities to raise awareness of the SDGs in education centres in the Autonomous Communities

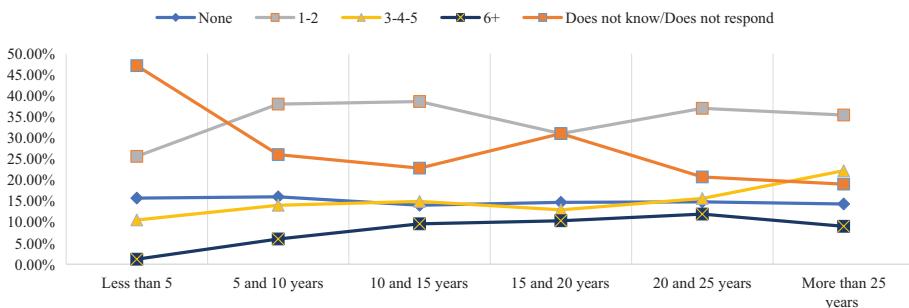


Source: Compiled by author

With regard to the number of years they have been teaching, the distribution is as follows: less than five years (20.8%; $n = 172$), between 5 and 10 years (12.1%; $n = 100$), between 10 and 15 years (13.8%; $n = 114$), between 15 and 20 years (14%; $n = 116$), between 20 and 25 years (16.3%; $n = 135$) and more than 25 years (22.9%; $n = 189$). It is confirmed that the

development of awareness-raising and sensitisation projects and activities by the school correlates with years of teaching ($p = .001$), specifically with those who have been teaching for less than five years (46.5%; $n = 80$) and those who have been teaching for more than 25 years (69.3%; $n = 131$). In addition, there is a significant relationship between the number of projects and activities the school has implemented ($p = .000$) and the years of teaching (FIGURE III).

FIGURE III. Distribution of the number of projects developed in the educational centres and the years of teaching practice



Source: Compiled by author

Regarding the development of curriculum competences and the degree of knowledge of the 2030 Agenda and the SDGs, there is a significant relationship ($p = < .05$) (TABLE III). Teachers who have a good or in-depth knowledge of the 2030 Agenda and the SDGs work on the following curriculum competences within the framework of the sustainable development approach: linguistic communication (14%, $n = 116$; 9.8%, $n = 81$ respectively), mathematics and basic competences in science and technology (13%, $n = 107$; 9.92%, $n = 82$ correspondingly), digital (15%, $n = 123$; 10.6%, $n = 88$); learning to learn (19%, $n = 156$; 14%, $n = 115$); social and civic (25.5%, $n = 211$; 18%, $n = 150$); sense of initiative and entrepreneurship (15.4%, $n = 128$; 10.7%; $n = 89$), and cultural awareness and expressions (16.3%, $n = 135$; 11.2%, $n = 93$).

TABLE III. Significance relationship between the level of knowledge of the 2030 Agenda, the SDGs and the curriculum competences

COMPETENCES	DEGREE OF KNOWLEDGE	
	Agenda 2030 (χ^2) Sig.	SDG (χ^2) Sig.
Linguistic communication	(7,750) .101	(17,878) .001
Mathematics and basic science and technology	(16,916) .002	(17, 486).002
Digital	(25,860) .000	(29, 402) .000
Learning to learn	(13,341) .010	(26,331) .000
Social and civic	(14,348) .006	(24,002) .000
Sense of initiative and entrepreneurship	(12,931) .012	(21,416) .000
Cultural awareness and expression	(12,229) .016	(16,468) .002

Source: Compiled by author

The results obtained show that the higher the level of knowledge of the 2030 Agenda and the SDGs by secondary school teachers, the greater the awareness and sensitisation to promote sustainable development in the classroom.

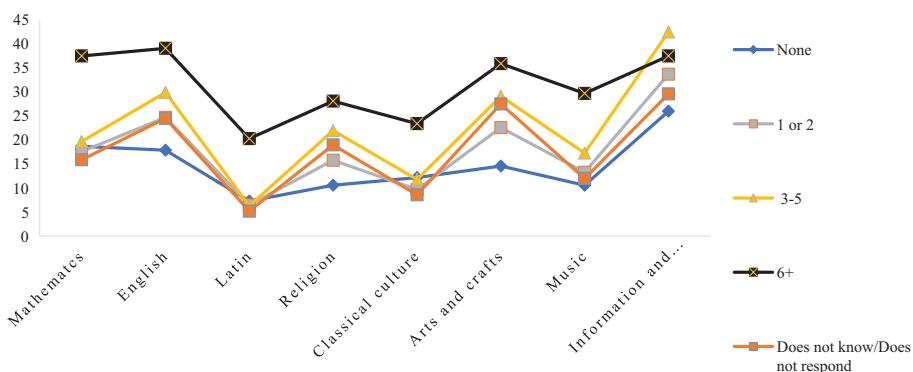
With regard to H2, developing projects and activities to raise awareness of the importance of the SDGs shows significant relationships ($p = < .05$) with linguistic competence ($p = .036$), learning to learn ($p = .005$), and social and civic competences ($p = .003$).

With regard to the variable Secondary Education subjects, significant relationships are detected between the degree of knowledge of the SDGs and the subjects of Economics ($p = .023$) (15%; $n = 124$), Initiation to Entrepreneurship and Business Activity ($p = .049$) (17.8%; $n = 101$), Ethical Values ($p = .002$) (17.4%; $n = 144$), and Scientific Culture ($p = .008$) (13.4%; $n = 111$), since the higher the degree of knowledge, the greater the implementation in the aforementioned subjects. There are no significant relationships between the subjects and the level of knowledge of the 2030 Agenda.

However, there are significant relationships ($p = < .05$) between the number of projects and activities to raise awareness of the importance of the SDGs that have been implemented in the school and the subjects Mathematics ($p = .003$), English ($p = .002$), Religion ($p = .020$), Plastic, Visual and Audiovisual Education ($p = .007$), and

Information and Communication Technology ($p = .045$) (FIGURE IV). The results show that the more projects are implemented in the school, the more they are incorporated into the subjects, especially in Mathematics, English, Art Education, and Information and Communication Technology.

FIGURE IV. Development of projects and activities and incorporation into the secondary secondary school curriculum



Source: Compiled by author

H2 of this research is accepted, as 77.5% ($n = 640$) of teachers reported having little, very little or no knowledge about the 2030 Agenda and 69.1% ($n = 571$) about the SDGs; compared to 22.5% ($n = 186$) and 30.9% ($n = 255$), respectively, who claim to know them well or in depth, showing that there are barriers to their embedding in the teaching-learning process towards the sustainability approach. Contradictorily, 78% ($n = 644$) of teachers' responses indicate that training is not a barrier; however, the analysis of the data reveals that there is a significant relationship ($p = < .05$) between the need for training by the teaching staff, the degree of knowledge of the 2030 Agenda ($p = .003$) and the SDGs ($p = .000$) (TABLE IV) to move towards curricular sustainability in the classroom.

TABLE IV. Difficulties in implementing sustainable development in the classroom depending on teachers' awareness of Agenda 2030 and SDGs

DIFFICULTIES	DEGREE OF KNOWLEDGE	
	Agenda 2030 (χ^2) Sig.	SDG (χ^2) Sig.
Excessive teaching load	(5,984) .200	(5,264) .261
Not explicit in the curriculum	(3,720) .445	(3,479) .481
The school's organisational documents do not provide for it.	(2,217) .696	(1,141) .888
The teaching staff needs training	(16,084) .003	(20,279) .000
The management team does not consider it necessary	(2,545) .637	(1,885) .757
The educational community is not aware of it	(1,285) .864	(2,391) .664
I don't think it would entail any difficulty	(7,116) .130	(12,835) .012

Source: Compiled by author

In this sense, and in response to the second objective, the barriers detected to implement the 2030 Agenda and the SDGs in secondary classrooms seem to fall on the direct actions of teachers, which focus on the organisational management of time for curriculum innovation, with an excessive teaching load (25.9%; $n = 214$) and their own training (22%; $n = 182$) being understood as an obstacle to entering into processes of curriculum innovation in their own teaching practice.

Finally, there are significant differences between the difficulties in developing the 2030 Agenda and the SDGs in the classroom, and the development of activities and projects to raise awareness of the importance of the SDGs. In this sense, there is a significant relationship that could lead to difficulties in their implementation in 0.2% ($n = 1$) of teachers because the documents of the organisational centre do not contemplate it ($p = .009$) and in 8.4% ($n = 42$) because there is no culture of awareness in the educational community ($p = .001$).

Discussion

The transition towards a society based on sustainable development involves a long process of social learning (Fernández & Gutiérrez, 2014) and the school can be an essential vector in this construction. Following authors such as Bonil et al. (2012), any sustainability process to be

introduced in schools must respond to the integration of sustainability processes from a structural dimension that is reflected in policies, in the cultures of schools and in educational practices; that is committed to rethinking the teaching-learning process, facilitating a continuous dialogue with the context to find new ways of understanding the balance between human beings, nature and society. Our study, endorsing the bibliography mentioned, shows that the development of awareness-raising and sensitisation projects and activities by the centre is configured as a facilitating variable for change in educational practices.

In the light of the evidence found regarding possible barriers to the implementation of actions for sustainability, centred on the excess of teaching hours and the need for training, in line with the results of Pramling and Kaga, (2008), it is necessary to incorporate teaching innovation as a key process for the sustainability of schools and curricula in Compulsory Secondary Education through processes of reflection and construction of meanings that enable teachers to empower themselves to guide their professional knowledge (Martínez, 2022), while promoting a process of shared leadership with the educational community, as indicated by Inoue et al. (2017), for the construction of conceptual, procedural and attitudinal content from the interdisciplinary complexity required to understand and respond to the current challenges of the planet (Albareda et al., 2017; Lasen et al., 2017). Change and quality in education “do not come from a clear and precise design, whether in the formulation of objectives or competencies, but from the demand for the empowerment and professional autonomy of teachers” (Martínez, 2022, p.139), which is why teaching innovation from the classroom-school context could be an enhancer for the know-how of teachers towards a collective formulation of new goals and strategies aimed at building education from the complexity and uncertainty of the processes involved in the current model of human development.

The limited knowledge of the SDGs and the 2030 Agenda itself among the teachers interviewed, in line with the research carried out by Inoue et al. (2017), and the small number of projects of this nature in terms of years of professional practice, leads to the need to address intergenerational training processes. Through contextualised training –initial and ongoing– that allows us to respond to collaborative needs (Imbernón, 2022), we will move towards the creation of new school environments that foster dialogic processes to address local and global eco-social challenges and challenges within formative contexts (Novoa & Alvim, 2022).

We believe, like authors such as Sureda-Negre et al. (2013), that one of the causes of the low level of knowledge among teachers and the low number of educational projects on sustainability in schools is the fragmented and unbalanced approach to the term that was included in Royal Decree 1631/2006, which develops a competency framework based on the so-called key competences for lifelong learning (Unión Europea, 2007), without including competences in sustainability. Perhaps the new decree on minimum teachings, Royal Decree 217/2022 of 29 March, which establishes the organisation and minimum teachings of Compulsory Secondary Education, will enable greater development of the key competences in current education systems (Patta-Tomas & Murga-Menoyo, 2020) by making their incorporation explicit, albeit in a transversal manner, in the training that students must attain at the end of this stage. Even so, we consider that the absence in the current regulatory framework of what are known as sustainability competences –incorporated as key, specific or transversal competences– could be a barrier to the implementation of actions leading to curricular sustainability.

Conclusions

Incorporating the sustainability approach in secondary education classrooms requires identifying possible barriers that may be present in teaching practices in order to provide keys that can “contribute to the education of a responsible citizenry, prepared to participate in decision-making and to address the serious socio-environmental problems facing humanity by adopting informed measures” (Vilches & Gil, 2012, p.25).

The results of the research presented here identify some challenges that need to be addressed not only by the education community, but also through policy frameworks. These correspond to the need to:

- Incorporate key competences in the field of sustainability –systemic thinking, anticipatory thinking, strategic thinking, collaboration, critical thinking, self-awareness and problem solving– into teacher training (UNESCO, 2017), so that they can transfer them to their practices. It is therefore necessary to focus on contextualised initial and ongoing teacher training that equitably and fairly guarantees the training of a citizenry that is sensitive and committed to current and future environmental and socio-economic needs and demands.

- Reinforce the training and professional development of secondary school teachers who have been teaching for between 5 and 25 years, so that they receive training in the sustainability approach, since in the light of the results obtained, they are less involved in the development of projects and activities linked to the sustainability approach. Incorporate this group into reflective theoretical-practical training processes (Martínez, 2022) that are constructive of meanings, based on the problematisation derived from current development models and leading to the construction of a teaching identity aligned with the sustainability approach, as they are a fundamental vector of change in the short and medium term and therefore of impact on future generations.
- It is considered necessary for schools to develop interdisciplinary projects or other types of activities linked to the sustainability approach, providing spaces during school hours to “deliberate on how to do things together, to make decisions collegially and to train in choice and responsibility” (Meirieu, 2022, p.187), for teachers to acquire not only a greater awareness and knowledge, but to embed it in their teaching practice with a holistic approach, integrating the various purposes of education: social, economic, cultural, ethical, environmental or otherwise. Only in this way can we move towards a systemic transformation of the current approach of conceiving education as watertight, non-interconnected disciplines.
- Educational communities need to be made aware of the importance of education for sustainable development, driven by the teaching staff, in order to achieve the multiplier effect of education.

To reduce them, some key points are proposed that can be taken into account. These are:

- Promoting educational policies that ensure the incorporation of the sustainability approach in the teaching-learning processes in order to respond to social requirements, promote values and build a more united world (Meirieu, 2022).
- Teacher training to promote awareness and sensitisation, while providing teaching tools to promote a transformative approach in the different subjects from a holistic and interdisciplinary approach

(Alcalá et al., 2020), as it has been ratified, as a key vector, that the degree of knowledge of the 2030 Agenda and the SDGs by secondary school teachers is a facilitator for incorporating the sustainability approach in the classroom.

- Having management and teaching teams in schools that exercise leadership in favour of the sustainability approach, with a proactive attitude towards change, moving towards more humanistic educational models, sharing a clear vision of a sustainable future, and being able to foster innovative processes that promote curricular sustainability.
- Considering the advisability of creating, strengthening and reinforcing inter-school collaboration networks with a vocation to grow and project themselves towards state and international networks (Fernández & Gutiérrez, 2014).

Limitations and prospects

The study presented here, although it explores the level of knowledge of secondary school teachers on sustainability and identifies some challenges that need to be faced in order to promote curricular sustainability in their classrooms, may be limited by the fact that the questionnaire does not include a question aimed at identifying the teaching specialty, since the current challenges facing the planet require a holistic and interdisciplinary approach, and not just a reference to subjects such as Biology and Geology, and/or Physics and Chemistry with a high degree of awareness and sensitivity to environmental issues, as is reflected in the study in the question asked about which subjects would incorporate the sustainability approach.

Royal Decree 217/2022, of 29 March (BOE, 2022), incorporates education for sustainability as a pedagogical principle to be developed in a cross-cutting manner; as well as specifying for the ESO exit profile that the key competences should be linked to the main global challenges of the 21st century (UNESCO, 2019), as well as the Sustainable Development Goals of the 2030 Agenda (UN, 2015).

This new pedagogical principle, which is incorporated into the current regulatory framework for ESO, is based on the need for knowledge, skills and attitudes to be addressed from a holistic and systemic approach to the different areas, fields and subjects that make up the curriculum.

This questioning leads us to consider that it would be interesting to investigate curricular sustainability in secondary classrooms according to the teaching specialty, in order to detect the real challenges, according to specialty, to which teachers are exposed in their teaching practice in the face of the challenge of the approach of education for sustainable development, and to propose possible keys that contribute to improving the teaching-learning process.

Do secondary school teachers understand sustainability as a holistic, systemic and therefore interdisciplinary construct that must modify teaching practices? Does the specialisation of secondary school teachers mediate the incorporation of the sustainability approach? These questions lead us to think that, although a high percentage of the teachers participating in the study may not have considered training as an obstacle to curricular sustainability, there is still a need, in the face of this new pedagogical challenge, for an update in the ongoing training of teachers.

Bibliographical references

Albareda, S., Fernández, M., Mallarach, J.M., & Vidal, S. (2017). Barreras para la sostenibilidad integral en la Universidad. *Revista Iberoamericana de Educación*, 73, 253-272. <https://doi.org/10.35362/rie730301>

Alcalá, M.J., Santos, M.J., Leiva, J.J., & Matas, A. (2020). Sostenibilidad curricular: una mirada desde las aportaciones del profesorado de la Universidad de Málaga. *Revista Internacional de Educación para la Justicia Social*, 9(2), 309-326. <https://doi.org/10.15366/riejs2020.9.2.015>

Álvarez-García, O., Sureda-Negre, J., & Comas-Forgas, R. (2018). Evaluación de las competencias ambientales del profesorado de primaria en formación inicial: estudio de caso. *Enseñanza de las Ciencias*, 36(1), 117-141. <https://doi.org/10.5565/rev/ensciencias.2338>

Azcárate, P., González-Aragón, C., Guerrero-Bey, A., & Cardeñoso, J.M. (2016). Análisis de la presencia de la sostenibilidad en los planes de estudios de los grados: un instrumento para su análisis. *Educar*, 52(2), 263-284. <http://dx.doi.org/10.5565/rev/educar.745>

Aznar, P., Ull, M.A., Martínez-Agut, M.P., & Piñero, A. (2017). La evaluación de la formación de formadores. Un catalizador en el proceso de cambio curricular hacia la sostenibilidad. *Revista Iberoamericana de Educación*, 73, 225-252. <https://doi.org/10.35362/rie730300>

Bermúdez, G., & Lía, A. (2008). La educación ambiental y la ecología como ciencia. Una discusión necesaria para la enseñanza. *Revista Electrónica de Enseñanza de las Ciencias*, 7 (2), 275-297. http://reec.uvigo.es/volumenes/volumen7/ART1_Vol7_N2.pdf

Bonil, J., Calafell, G., Granados, J., Junyent, M., & Tarín R.M. (2012). Un modelo formativo para avanzar en la ambientalización curricular. *Profesorado. Revista de currículum y formación del profesorado*, 16(2), 145-163. <https://recyt.fecyt.es/index.php/profesorado/article/view/43723>

Brundiers, K., Barth, M., Cebrián, G., Cohen, M., Diaz, L., Doucette-Remington, S., Dripps, W., Habron, G., Harré, N., Jarchow, M., Losch, K., Michel, J., Mochizuki, Y., Rieckmann, M., Parnell, R., Walker, P., & Zint, M. (2021). Key competencies in sustainability in higher education-toward an agreed-upon reference framework. *Sustainability Science*, 16, 13-29. <https://doi.org/10.1007/s11625-020-00838-2>

Calero, M., Mayoral, O., Ull, Á., & Vilches, A. (2019). La educación para la sostenibilidad en la formación del profesorado de ciencias experimentales en Secundaria. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 37(1), 157-75. <https://doi.org/10.5565/rev/ensciencias.2605>

Fernández, R., & Casado, O. (2017). "Somos uno". La educación como motor de cambio social. *Revista de Infancia, Educación y Aprendizaje*, 3(2), 486-491. <https://revistas.uv.cl/index.php/IEYA/article/view/768/744>

Fernández, M.A., & Gutiérrez, J.M. (2014). *La educación hacia la sostenibilidad en la CAPV. Contribución de la educación ambiental a la difusión de la cultura de la sostenibilidad*. Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia.

Filho, W.L., Brandili, L.L., Castro, P.C., & Newman, J. (2017). *Handbook of Theory and Practice of Sustainable Development in Higher Education*. Springer

Fundación Benéfico-Social Hogar del empleado (FUHEM). (2018). *Educar para la transformación ecosocial. Orientaciones para la incorporación de la dimensión ecosocial al currículo*. FUHEM ecosocial

Gil,D., Vilches,A., Toscano,J.C.,& Macías,Ó.(2006).Década de la Educación para un Futuro sostenible (2005-2014): un punto de inflexión necesario en la atención a la situación del planeta. *Revista Iberoamericana de Educación*, 40(1), 125-178. <https://rieoei.org/historico/documentos/rie40a06.pdf>

Imbernón, F. (2022). Políticas y prácticas educativas en la formación del profesorado. En En C. Rodríguez y F. Imbernón (Coords.), *De las políticas educativas a las prácticas escolares* (pp. 283-296). Morata

Inoue, M., O'Gorman, L., Davis, J., & Ji, O. (2017). An international comparison of early childhood educator's understandings and practices in education for sustainability in Japan, Australian, and Korea. *International Journal of Early Childhood*, 49, 353-373. <https://doi.org/10.1007/s13158-017-0205-5>

Lansen, M., Skamp, K., & Simoncini, K. (2017). Teacher perceptions and self-reported practices of education for sustainability in the early years of primary school: An Australian case study. *International Journal of Early Childhood*, 39, 391-410. <https://doi.org/10.1007/s13158-017-0200-x>

Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boletín Oficial del Estado (BOE)*, núm. 340, de 30 de diciembre de 2020, 122868-122953.

Martínez, J. (2022). Políticas del currículum: complejidad, participación, culturas y hegemonía. En C. Rodríguez y F. Imbernón (Coords.), *De las políticas educativas a las prácticas escolares* (pp. 133-150). Morata

McKeown-Ice, R. (2000). Environmental education in the United States: A survey of preservice teacher education programs. *Journal of Environmental Education*, 32(1), 4-11. <https://doi.org/10.1080/00958960009598666>

Meirieu, P. (2022). Fines y modos de nuestras instituciones educativas. En C. Rodríguez y F. Imbernón (Coords.), *De las políticas educativas a las prácticas escolares* (pp.187-198). Morata

Murga-Menoyo, M.Á., & Novo, M. (2014). Sostenibilizar el currículum. La Carta de la Tierra como marco teórico. *EDETANIA*, 46, 163-179. <https://revistas.ucv.es/edetania/index.php/Edetania/article/view/166/141>

Murga-Menoyo, M.Á., & Novo, M. (2017). Sostenibilidad, desarrollo 'global' y ciudadanía planetaria. Referentes de una pedagogía para el desarrollo sostenible. *Teoría de la Educación. Revista Interuniversitaria*, 29(1), 55-78. <https://doi.org/10.14201/teoredu2915579>

Murga-Menoyo, M. Á. (2021). La educación en el Antropoceno. Posibilismo versus utopía. *Teoría de la Educación. Revista Interuniversitaria*, 33(2), 107-128. <https://doi.org/10.14201/teri.25375>

Novoa, A., & Alvim, Y. (2022). Los profesores después de la pandemia. En C. Rodríguez y F. Imbernón (Coords.), *De las políticas educativas a las prácticas escolares* (pp. 29-48). Morata

Olaskoaga-Larrauri, J., Guerenabarrena-Cortazar, L., & Cilleruelo-Carrasco, E. (2021). Academic staff attitudes and barriers to integrating sustainability in the curriculum at Spanish universities. *Culture and Education*, 33(2), 373-396. <https://doi.org/10.1080/11356405.2021.1905957>

Orden ECI/3858/2007, de 27 de diciembre, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de las profesiones de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas. *Boletín Oficial del Estado (BOE)*, núm. 312, de 29 de diciembre de 2007, 1-6.

Patta-Tomás, M., & Murga-Menoyo, M.A. (2020). El marco curricular de la Educación Secundaria Obligatoria: posibilidades para la formación de competencias en sostenibilidad. *Revista internacional de comunicación y desarrollo*, 3(13), 90-109. <https://doi.org/10.1080/11356405.2021.1905957>

Pramling, I., & Kaga, Y. (2008). *The contribution of early childhood education to a sustainable society*. UNESCO

Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria. *Boletín Oficial del Estado (BOE)*, núm. 76, de 30 de marzo de 2022, 41571-41789.

Reyes, M.M. (2010). La formación del profesorado Motor de cambio en la escuela del siglo XXI. *Revista de Educación Inclusiva*, 3(3), 89-102. <https://revistaeducacioninclusiva.es/index.php/REI/article/view/175/169>

Scharenberg, K., Waltner, E.M., Mischo, C., & Rieß, W. (2021). Development of students' Sustainability Competencies: Do teachers make a difference? *Sustainability*, 13(22), 12594. <https://doi.org/10.3390/su132212594>

Southwell, M. (2013). *La escuela ante nuevos desafíos: participación, ciudadanía y nuevas alfabetizaciones*. Santillana

Sureda-Negre, J., Catalán-Fernández, A., Álvarez-García, O., & Comas-Forgas, R. (2013). El concepto de “desarrollo sostenible” en la regulación del currículum de la Educación Secundaria Obligatoria en España. *Estudios Pedagógicos*, 39(1), 253-267. <http://dx.doi.org/10.4067/S0718-07052013000100015>

Unión Europea. (2007). *Competencias clave para el Aprendizaje Permanente. Un marco de Referencia Europeo*. <https://n9.cl/483y>

United Nations (UN). (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf

United Nations Educational, Scientific and Cultural Organization (UNESCO). (s.f.). *UN Decade of ESD*. <https://en.unesco.org/themes/education-sustainable-development/what-is-esd/un-decade-of-esd>

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2017). *Education for Sustainable Development Goals: learning objectives*. <https://unesdoc.unesco.org/ark:/48223/pf0000247444>

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2019). *Reconceptualizing and Repositioning Curriculum in the 21st Century: A Global Paradigm Shift*. <http://www.ibe.unesco.org/es/news/reconceptualizing-and-repositioning-curriculum-21st-c-global-paradigm-shift>

Valderrama, R., Alcántara, L., Sánchez, F., Caballero, D., Gil, D., Vidal, S. & Miñano, R. (2020). ¿Forma en sostenibilidad el sistema universitario español? Visión del alumnado de cuatro universidades. *Educación XXI*, 23(1), 221-245. <https://doi.org/10.5944/educxx1.23420>

Vilches, A., & Gil, D. (2012). La educación para la sostenibilidad en la universidad: el reto de la formación del profesorado. *Profesorado. Revista de Currículum y Formación del Profesorado*, 16(2), 26-43. <https://www.ugr.es/~recfpro/rev162ART3.pdf>

Vilches, A., & Gil, D. (2013). La ciencia de la sostenibilidad en la formación del profesorado de ciencias. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10(número extraordinario), 749-762. http://dx.doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2013.v10.iextra.17

Contact address: Belén Sáenz-Rico de Santiago. Universidad Complutense de Madrid, Facultad de Educación, Dpto. Estudios Educativos. Edificio – La Almudena. C/Rector Royo Villanova, 1. Ciudad Universitaria (28040, Madrid). E-mail: bsaenzri@edu.ucm.es

Investigating the efficacy of retrieval practice in university mathematics

Investigación de la eficacia del aprendizaje potenciado por recuperación en las matemáticas universitarias

<https://doi.org/10.4438/1988-592X-RE-2023-401-584>

Csaba Szabó

<https://orcid.org/0000-0003-4783-5411>

Eötvös Loránd University, Faculty of Science

MTA-ELTE Theory of Learning Mathematics Research Group

Csilla Zámbó

<https://orcid.org/0000-0002-7230-7023>

Eötvös Loránd University, Faculty of Primary and Pre-School Education

MTA-ELTE Theory of Learning Mathematics Research Group

Anna Muzsnay

<https://orcid.org/0000-0003-4045-0527>

University of Debrecen

MTA-ELTE Theory of Learning Mathematics Research Group

Janka Szeibert

<https://orcid.org/0000-0002-4267-9360>

Eötvös Loránd University, Faculty of Primary and Pre-School Education

MTA-ELTE Theory of Learning Mathematics Research Group

László Bernáth

<https://orcid.org/0000-0002-8314-6560>

Eötvös Loránd University, Faculty of Education and Psychology

MTA-ELTE Theory of Learning Mathematics Research Group

Abstract

Retrieving information from memory can strengthen one's memory of the retrieved information itself. The strategic use of retrieval to enhance memory and help long-term retention is known as test-enhanced learning or retrieval practice. Test-enhanced learning has been proven effective concerning different learning materials, but these experiments were primarily conducted in laboratory environments and focused mainly on memorization. Our aim was to explore the efficacy of test-enhanced learning used for teaching mathematics at university level. The experiment was carried out in classroom settings, concerning obligatory courses. The participants were six groups of undergraduate pre-service mathematics teachers. Three groups learned Number Theory using the testing effect, and the other three learned using traditional methods. The experimental and control groups learned the exact same information in the lecture and wrote the same final test. The experimental group performed significantly better than the control group, although their performance on the initial competence exams was significantly worse. The results indicate that test-enhanced learning has a significant advantage in solving complex mathematical problems. To examine the effect of differences in individual competence, we divided the students in both experimental and control groups into low-, middle-, and high-performing groups. The efficacy of test-enhanced learning was demonstrated in all the three performance levels. Regarding the three pairs of groups, members of the experimental group using test-enhanced learning performed better than those of the control group.

Keywords: testing effect, retrieval practice, mathematics, complex problems, individual differences.

Resumen

Recuperar información de la memoria puede reforzar el recuerdo de la propia información recuperada. (Este es el fenómeno llamado “efecto test” o “testing effect”.) El uso estratégico de la recuperación para mejorar la memoria y ayudar a la retención a largo plazo se conoce como aprendizaje potenciado por pruebas o aprendizaje potenciado por recuperación. El aprendizaje potenciado por pruebas ha demostrado su eficacia en relación con diferentes materiales de aprendizaje, pero estos experimentos se realizaron principalmente en entornos de laboratorio y se centraron sobre todo en la memorización. Nuestro objetivo era explorar la eficacia del aprendizaje potenciado por pruebas utilizado para la enseñanza de las matemáticas a nivel universitario. El experimento se llevó a cabo en entornos de aula, en relación con cursos obligatorios. Los participantes fueron seis grupos de profesores de matemáticas en formación. Tres grupos aprendieron Teoría de los Números utilizando el efecto test, y los otros tres aprendieron utilizando métodos tradicionales. Los grupos experimental y de

control aprendieron exactamente la misma información en la clase y realizaron el mismo examen final. El grupo experimental obtuvo un rendimiento significativamente mejor que el grupo de control, aunque su rendimiento en los exámenes de competencia inicial fue significativamente peor. Los resultados indican que el aprendizaje potenciado por los exámenes tiene una ventaja significativa en el aprendizaje de la resolución de problemas matemáticos complejos. Para examinar el efecto de las diferencias en la competencia individual, dividimos a los estudiantes de los grupos experimental y de control en grupos de rendimiento bajo, medio y alto. La eficacia del aprendizaje potenciado por los exámenes se demostró en los tres niveles de rendimiento. En cuanto a los tres pares de grupos, los miembros del grupo experimental que utilizó el aprendizaje potenciado por tests obtuvieron mejores resultados que los del grupo de control.

Palabras clave: efecto del test, aprendizaje potenciado por recuperación, matemáticas, problemas complejos, diferencias individuales.

Introduction - Theoretical background

Retrieving information from memory after an initial learning phase enhances long-term retention more than restudying the material, an advantage referred to as the testing effect (Roediger & Butler, 2011; Rowland, 2014). The testing effect has been demonstrated with various practice tests, materials, and age groups (Karpicke, 2017) including tertiary education (Butler, 2010). However, these experiments were mostly conducted in laboratory environments concerning the memorization of texts or words. There have been only a limited number of experiments in the field of mathematics, in real-life educational environments (Lyle & Crawford, 2011; Lyle, Hopkins et al., 2016; Fazio 2019; Lyle, Bego et al., 2020).

Testing enhances the effectiveness not only of word-for-word retention - as opposed to rereading - but also of the application of the newly acquired knowledge. Smith and Karpicke (2014) have shown that groups studying with tests performed better than the control group not only in word-for-word retention tasks but also in tasks demanding the synthesis of the information within the given text. Furthermore, knowledge acquired with the use of retrieval practice was not only shown to be more applicable within the given subject of the text than rereading-based knowledge but also more easily transferable to other areas (Butler, 2010; van Eersel et al., 2016).

The strong evidence for a direct effect of testing suggests that retrieval practice may be regarded as one of the most effective learning techniques. (Karpicke & Blunt, 2011; Larsen et al., 2013; Dunlosky et al., 2013; Donoghue & Hattie, 2021). However, there are certain areas where the demonstration of the testing effect has either failed or produced contradicting results.

A potential impediment concerns the results of Khanna (2015) based on a study conducted within an introductory psychology course. Students were placed into “ungraded quiz,” “graded quiz,” and “no quiz” groups, and the first two were given six surprise tests in the semester. The results show that the “ungraded” group performed better than the other two groups, and there was no difference in performance between the “graded” and “no quiz” groups. Khanna’s explanation for this result is that higher levels of anxiety can eliminate the testing effect on the “graded” group. However, these conclusions contradict the findings of Agarwal et al., (2012) that test anxiety decreases in students studying with frequent testing and the results of Tse & Pu (2012) that demonstrated the efficiency of testing effect in case of people both with low and high anxiety. One possible solution to this contradiction is that anxiety only impairs study performance in the case of weak intrinsic motivation, and, if backed up by higher intrinsic motivation, anxiety may, in fact, serve to improve performance (Wang et al., 2015). The results of Emmerdinger and Kuhbandner (2019) can give another solution to the contradiction. They found that the testing effect appears independently of the emotional state (negative, neutral, or positive) of the participants.

Concerning the form of testing, both short-answer and multiple-choice tests are more effective ways of learning than rereading (Kang et al, McDermott, and Roediger, 2007). The efficacy of testing may vary, depending on the presence of feedback. If feedback is included, short-answer questions are more beneficial; otherwise, multiple-choice tests are more effective (Kang et al., 2007).

Another contradiction concerns the role of individual differences in test-enhanced learning. Orr and Foster (2017) conducted their examination within a biology course in which students had the option of participating in tests administered periodically throughout. Those who systematically took part in the tests performed better in the final exam than those who did not. Furthermore, most important from our viewpoint is that this advantage was observed in students with above-

average, average, and below-average skills alike. By contrast, the results of Carpenter et al. (2016), also within a biology course, show that test-enhanced learning was only effective in students with above-average skills and that no enhancement was observed in students with average and below-average skills. This greatly impedes the application of test-enhanced learning in a classroom setting. Nevertheless, the results of Carpenter et al. (2016) contradict the assertion of Brewer and Unsworth (2012) that individuals with lower general-fluid intelligence (Gf-I) profited more from test-based studying than individuals with higher Gf-I and that test-enhanced learning could not be observed at all concerning the highest-level Gf-I individuals. Furthermore, Balota et al. (2006) demonstrated the benefits of testing effect among people having dementia of Alzheimer's type. This result suggests that testing is beneficial for those with high cognitive abilities, for those with average or below average abilities, and even for memory-impaired people.

The third factor that has produced contradicting results concerns the role of the level of complexity of the object of study. Van Gog and Sweller (2015) argue that the testing effect can only be observed when there is no interaction between the items to be learned, for instance when learning the vocabulary of a foreign language; in more complex subject materials, it either diminishes or disappears completely. (However, see Karpicke and Aue (2015) for theoretical counterarguments). Leahy et al. (2015) observed the testing effect related to complex study material upon immediate retention testing, and they did not detect any effect at all following one week's delay. A similarly negative result was arrived at by Tran et al. (2015), whose "revision" and "testing" groups had to learn consecutively appearing sentences describing various scenarios. In the final test, although the retention of the individual sentences showed the testing effect, there was no difference in performance between the two groups in terms of drawing conclusions based on the content of the sentences. In other words, they found that, in complex tasks requiring deductive thinking, the testing effect disappears. Eglington and Kang (2018) repeated the experiment of Tran et al. (2015) with one modification (the sentences were shown on the monitor all at once, not one by one); their results demonstrated the benefits of the testing effect on this deductive task.

In the study of Peterson and Wissmann (2018), the retrieval effect had no advantage compared to restudying in the case of solving complex problems requiring analogical thinking. Despite these results, Wong et al (2019) and Hostetter et al. (2019) proved the benefits of retrieval learning in analogical problem-solving. Most likely, in the case of Peterson and Wissmann

(2018), the inefficacy of the retrieval effect can be explained not (or not only) by the requirement of analogical thinking, but maybe (also) by the complexity of the problems or some other elements of the experimental design. According to the abovementioned results, in tasks demanding complex or deductive thinking, the advantage of test-enhanced learning over rereading learning is unclear. Furthermore, as the results of Carpenter et al. (2016) and Brewer and Unsworth (2012) have shown, the role of individual differences in competence is also ambiguous.

Mathematical problems require developed deductive and problem-solving skills, and the problems themselves are quite complex. Developing problem-solving skills in mathematics requires the application of procedures and deep conceptual understanding, not only memorization. Although there have been only a few investigations about the testing effect on mathematical problem-solving in a real school environment, recent studies suggest that the intensive use of retrieval practice may be an effective way of learning (Lyle and Crawford, 2011; Fazio 2019; Lyle et al., 2016; Lyle et al., 2020). The paper of Avvisati and Borgonovi (2020) concerns problem-solving in mathematics. Although the environment is not a real educational environment in the sense that they measured the effect of a single test practice, their large sample investigation is relevant for us since it uses educational material. They demonstrate that the number of mathematical problems in the first test had a small positive effect on the average mathematics performance on the second test. In the experiment of Yeo and Fazio (2019), the efficacy of retrieval practice and worked examples for different learning goals were examined. The optimal learning strategy depended on the retention interval and the nature of the materials. Repeated testing was more effective than repeated studying after a 1-week delay when students' goal was to remember the text of a worked example. On the other hand, they found that learning a novel maths procedure and measuring performance immediately, repeated studying was more optimal than repeated testing, regardless of the nature of the materials. Finally, the study of Lyle et al. (2011) is the most relevant study for our research. They incorporated retrieval practice into a course on statistics for psychology by adding a brief retrieval exercise for some essential lecture material at the end of every lecture. This method significantly and substantially increased exam scores. Students liked the retrieval practice and believed it was helpful. In Lyle's later studies, spaced retrieval practice was investigated in the precalculus

course. Their results have encouraged us to explore further aspects of the testing effect in university mathematics classes (Lyle et al. 2016, 2020).

When applying retrieval practice in a real school setting, two important questions arise: the form of testing and the differences in individuals' mathematical competencies. Concerning the form of testing, both short-answer and multiple-choice tests are more effective ways of learning than rereading (Kang et al. 2007). The efficacy of testing may vary, depending on the presence of feedback concerning the practice test. If feedback is included, short-answer questions are more beneficial; otherwise, multiple-choice tests are more effective (Kang et al., 2007). Also, the question of differences in individuals' mathematical competencies can be a central one. We will further discuss this aspect in the next chapter as its investigation was part of the aim of our study.

The aim of the study

In this study, we aimed to implement retrieval practice in university education and investigate its efficacy in learning higher mathematics. When applying retrieval practice in a real school environment in mathematics, it is not evident which method to use and how to put it into practice. We must pay attention to the fact that retrieving must take place within 24 hours, there should be no copying, no cheating, it should not take a lot of time, and students should be involved in it. The form of the testing and the type of questions have to be considered, as well.

Also, our goal is to examine if the efficacy of retrieval practice depends on the individuals' mathematical competence. In case of university maths education, this is a particularly important aspect, where differences among the entering students are usually enormous. For this reason, our study examines an actual educational environment and regularly applied course material. It compares the performance of two groups of students in an Algebra and Number Theory course studying with either a traditional method or a test-based method (recalling study material on one occasion, immediately after learning it) considering the effects of differences in individual competence. To examine the differences in individual competence, we grouped students based not on their all-around performance during the whole course, as Carpenter et al. (2016) did, but on a competence level test taken in the initial class.

Method

The authors conducted a quasi-experimental study to figure out whether retrieval effect leads to improvement in mathematics achievement at tertiary level.

Sample

The participants of the experiment were all first-year mathematics students at [] University, comprising 114 persons in total, attending the Algebra and Number Theory course. During the analysis, we discounted the data relating to students who had previously taken the course, leaving 72 persons in all, 26 male and 46 female. Their ages were between 18 and 23.

Materials

The regular course materials for the Algebra and Number Theory lectures and problem-solving seminars were used based on the textbook by Niven et al. (1991): *An Introduction to the Theory of Numbers*, 5th ed.

Procedure and instruments

The course, which the students attended in six groups, consisted of one 60-minute lecture and one 90-minute problem-solving seminar per week for a total of 13 weeks. Each student completed a competence-level test at the initial class. Three of the six groups were randomly selected as the experimental group, the other three were the control group; 37 students were in the experimental group, and 35 were in the control group. The teachers of the control groups and the experimental groups were in pairs. Pairs were created according to teaching experience, with 1-1 experienced teacher, 1-1 demonstrator, and 1-1 doctoral student in the control and experimental groups.

The problem-solving seminars consisted of tasks based on the theoretical subject matter of the previous week's lecture, which were solved collectively with the help of a professor. The structure of each lesson for the

control group was the following: at the beginning of the class, they wrote a short test on the previous week's material (as it is traditional in the case of this subject). This was followed by the discussion of homework and the main part, which is problem-solving with the aid of the professors. In the experimental group, the structure was almost identical, the only difference was that there was no test at the beginning of the lesson, instead, they had a test at the end of the class (see Appendix A). Our method resembled to that of Lyle and Crawford (2011) in that students had to complete tasks similar to those encountered during the class. The end-of-class test consisted of 2 problems and members of the experimental group had to solve it individually, without any help (while members of the control group solved it with the professors, as with all other problems). The solutions of the problems of the end-of-class test were not discussed in the experimental groups, only if the students asked for it. By solving a problem students could gain 1-1 points. In case their solution was perfect, they got 1 point. If they made a little mistake, they could still get 0,5 points. Otherwise, they got 0 points. When revising the papers, we tried to be objective. The teachers of the course had a short conversation every week when they discussed how to correct the tests. After these conversations, students' tests were corrected by their teacher. Finally, the lecturer reviewed the corrections of all the papers. All students (both in the control and experimental group) had to reach at least 50% of the aggregate score of the seminar tests. This was the prerequisite for the final examination.

In the last problem-solving seminar, both groups completed a final test consisting of five problems (see Appendix B), the score of which determined their final grades.

The evaluation of the final test was the following: the perfect solution for each problem was awarded 6 points. Sub-scores for partial solutions, evaluation of the, most common errors and the general policies of correction were included in a detailed scoring guide, written by the lecturer. Students' tests were corrected by their teacher (with the help of the lecturer, if consultation was necessary), according to these guidelines. Finally, the lecturer reviewed the corrections of all the papers.

The following contains some analysis of the problems in the final test. The first problem is a typical example requiring procedural knowledge. Procedural (vs. conceptual) means that this kind of problem always can be solved using the same method. Plus, both the procedure and the calculations are rather easy. This problem can even be solved using only high

school knowledge. However, the problem is still complex. Although the general method can be practiced to a great extent, you need a trick in each step, and this trick depends on the numbers appearing in the problem.

The second problem can also be solved procedurally, but it is one step more complicated. The knowledge required to solve this problem is brand new university knowledge; there is no way to solve it with high school techniques. The procedure, as in the previous case, is always the same, but an extra formula must be applied beforehand. (The formula is known by everybody.) As in the first problem, with a lot of practice, students can be prepared for the midterm, but for long-term application, conceptual knowledge is also required in this case. As well as in the previous case, you can acquire the general method, but you still need a trick, which depends on the actual numbers of the problem. This adds a lot to the complexity of the problem. This calculation cannot be reversed even by the most modern, fastest computers. For example, bank cryptography is based on this method: our computer generates a code using this method and this way the code cannot be broken.

The third problem requires all possible abstract skills gained during study. The concept it uses is rather difficult and is strongly connected to the notion of *multiplicative order* from abstract algebra, which is by far the hardest concept in the material. The knowledge of this concept can be checked in several different ways, and each way is challenging. This was the last topic taught before the midterm; hence, there was a chance that students did not have enough time to conceptualize the notion. Additionally, we must mention that this concept was not practiced during the problem-solving seminars, only during the lecture.

The fourth problem is the most complex. The solution requires the application of several different strategies such that each must be selected from some lists of different strategies, each of which contains infinite strategies. Although the material taught and the form of the problem suggested the lists of strategies, it is not obvious which of the lists has to be applied. If one of the strategies was wrongly selected, participants had to start solving the problem from the beginning. The solution of this problem definitely requires a conceptual understanding of most topics. Furthermore, this function has several ways of formulation and no matter which one we choose we must work out how it can be interpreted in order to use it for the actual problem. When we have the interpretation, it is still complex how to use it because the formula consists of several components.

Problem 5 is always the most difficult problem with a solution that is easy to understand and very hard to find. For the solution, you need to see the global structure of the material and make deductions on the features of the items based on the structure. Then you need to decide on which items the deduction is true and then step back to the structure to find whether there exists such item or not.

Results

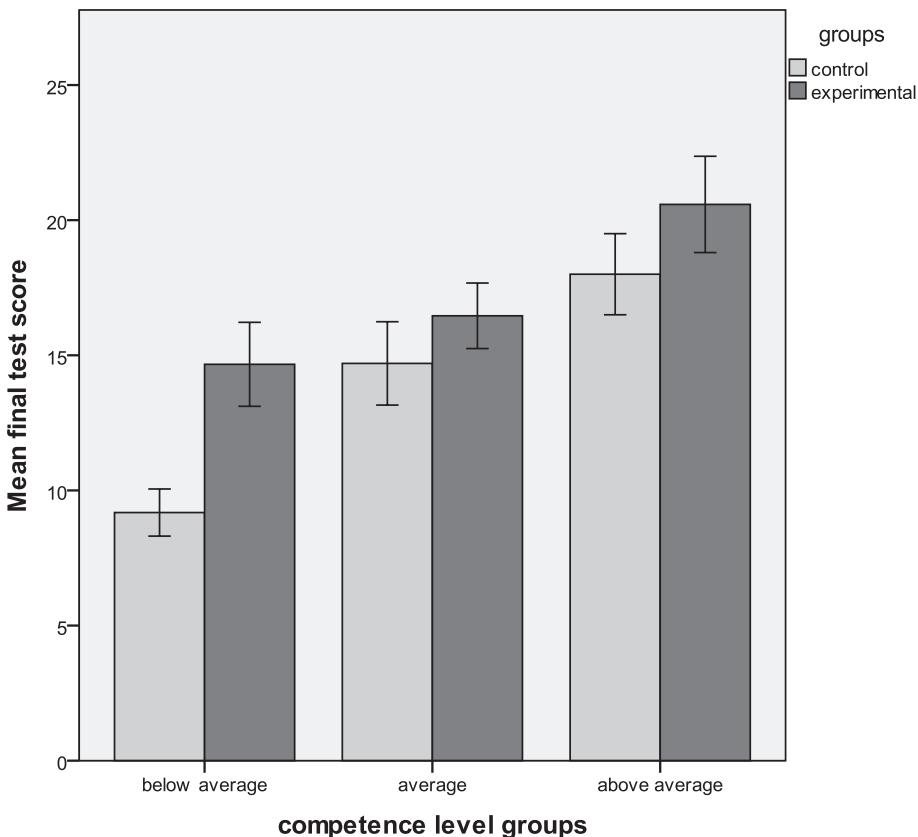
When analysing the results, we found that there is no difference among genders ($F(1,67)=0.29$ $p>0.05$ $\eta_p^2 = 0.004$), as well as no significant interaction ($F(1,67)=3.92$ $p>0.05$ $\eta_p^2 = 0.055$), so we can state that gender has no effect at all. Thus, in the following, we do not include this aspect in the analysis.

The score of the competence-level test was $M = 57.24$; $SD = 19.50$ in the experimental group and $M = 60.46$; $SD = 20.12$ in the control group. The score of the final exams was $M = 17.22$; $SD = 5.74$ in the experimental group and $M = 14.29$; $SD = 5.91$ in the control group, where the maximum score was 30.

We wanted to eliminate the effect of differences in prior knowledge on the scores of the final test, so we analysed the data using ANCOVA (Cohen, 1988, pp. 287-288). When comparing the experimental and control groups based on the final exam scores and controlling for their competence-level test scores, we found that the experimental group performed significantly better than the control group, $F(1,69) = 9.19$ $p<0.001$, $\eta_p^2 = 0.118$, despite the fact that their performance on the competence exams was significantly worse, $F(1,69)=32.79$ $p<0.001$ $\eta_p^2 = 0.322$

To examine the effect of differences in individual competence, we divided the students in both the experimental and control groups into below-average, average, and above-average groups based on their competence-level tests, placing those with $\pm 1/2$ SD around the mean into the “average” category, those with under $1/2$ SD of the mean into the “below average” category, and those with over $1/2$ SD of the mean placed into the “above average” category. The average scores of their final exams are shown in Figure I. below.

FIGURE I. The performance of below-average, average, and above-average students within the experimental and control groups. Error bars represent ± 1 SE.



Source: Compiled by author

The data were analysed using a 2×3 (experimental-control, below average-average-above average) ANCOVA. The experimental group performed significantly better than the control group $F(1,66) = 7.52$ $p < 0.001$, $\eta_p^2 = 0.102$; the difference of the groups based on competence is significant, $F(2,66) = 13.02$ $p < 0.001$, $\eta_p^2 = 0.283$; based on the Sidak correction for multiple comparisons, the performance of all three groups differed significantly from each other, and there was no significant interaction, $F(2,66) = 0.86$ $p > 0.05$, $\eta_p^2 = 0.026$; the testing effect was shown to be independent of individual competence.

Discussion

The purpose of this study was partly to examine if the advantages of test-enhanced learning over traditional learning techniques — problem-solving exercises — may be observed using the complex mathematical curriculum of an actual educational environment. This was important because the testing effect has not yet been demonstrated in relation to higher-level mathematics and an actual educational environment. Furthermore, test-enhanced learning was shown to lose its advantage when its object concerns deductive tasks and complex materials (Tran et al., 2015) or may result in weaker performance (Leahy et al., 2015). In contrast, our own results indicate that test-enhanced learning has a significant advantage in relation to solving complex mathematical problems. The learning process of the experimental group and the control group was tightly synchronised so that they get familiar with the same concepts, the same problems, and exactly the same tasks. The experimental group's performance showed significant improvement despite that this group had less mathematical understanding at the beginning of the semester. They reached better results in the final examination than the control group studying with traditional methods.

The other aim of our study was to investigate Khanna's (2015) idea: whether graded tasks in fact result in weaker performance. In their experiment, they were unable to demonstrate the benefits of test-enhanced learning over traditional revision practices. They posited that the negative results were due to anxiety caused by testing, which impedes performance. - students of their courses completed a 6-item questionnaire on their feelings about the inclusion of quizzes in the course. Other previous studies show the effectiveness of test-enhanced learning when the final test occurs under stress (such as in an exam) (Szöllősi et al., 2017). Multiple measures of stress levels were applied, such as anxiety tests like STAI and PANAM, and saliva sampling for measuring the cortisol level. The results favoured the ecological validity of retrieval-based learning. As we already mentioned in the introduction, anxiety backed up by higher intrinsic motivation may, in fact, serve to improve performance (Wang et al., 2015). Our results strengthen the positivity of the testing effect, in the sense that graded quizzes produced significantly better performance. We also examined the role of differences in individual competence as the existing literature on the subject was contradictory.

In Carpenter et al. (2016), the testing effect was only observed in above-average students while Orr and Foster (2017) identified it in below-average, average, and above-average individuals alike. In our opinion, the two experiments differ because, while Carpenter et al. (2016) grouped the participants according to competence only after the completion of the course based on their performance, Orr and Foster (2017) compared them based on the results of the first three tests. In our own study, we administered a preliminary test prior to the onset of the classes to gauge competence and thereby determine the below-average, average, or above-average status of participants. Our results show, similarly to those of Orr and Foster (2017), that the testing effect appears independently of individual mathematical competence.

One limitation within our study is that it does not uncover what other possible individual competencies may influence the effect of test-enhanced learning past one's mathematical competence—for example, Gf-I—as this question was not examined. Another limitation is that we did not monitor changes in test anxiety within the experimental and control groups, which might have been relevant in finding out whether decreased anxiety among the control group taking weekly tests could have been the (or a) cause for their better performance during the final examination. In order to reduce the teachers' effect, we chose “teacher-pairs” for the experiment, according to their teaching experience.

The question of finding the best way to test students with different mathematical abilities is an open and interesting one. We suspect that mathematical ability and how the test is designed (difficulty and form of questions) are somehow related to the rate of students' progress. In our experiments, we found that the kind of retrieval practice we applied can be an effective way of learning higher mathematics that professors can implement in their lessons to enhance students' performances and is beneficial for students with either low, average, or high mathematical competence.

Bibliographic References

Agarwal, P. K., Bain, P. M., & Chamberlain, R. W. (2012). The value of applied research: retrieval practice improves classroom learning and recommendations from a teacher, a principal, and a scientist. *Educational Psychology Review*, 24(3), 437–448. <https://doi.org/10.1007/s10648-012-9210-2>

Avvisati, F., & Borgonovi, F. (2020). Learning mathematics problem solving through test practice: A randomized field experiment on a global scale. *Educational Psychology Review*, 32(3), 791–814. <https://doi.org/10.1007/s10648-020-09520-6>

Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger, H.L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychol. Aging*, 21(1), 19–31. <https://doi.org/10.1037/0882-7974.21.1.19>

Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66, 407–415. <https://doi.org/10.1016/j.jml.2011.12.009>

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of experimental psychology. Learning, memory, and cognition*, 36(5), 1118-1133. <https://doi.org/10.1037/a0019902>

Carpenter, S. K., Lund, T. J. S., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A classroom study on the relationship between student achievement and retrieval-enhanced learning. *Educational Psychology Review*, 28(2), 353–375. <https://doi.org/10.1037/xap0000145>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Donoghue, G. M., & Hattie, J. A. C. (2021). A meta-analysis of ten learning techniques. *Frontiers in Education*, 6:581216. <https://doi.org/10.3389/feduc.2021.581216>

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1) 4–58. <https://doi.org/10.1177/1529100612453266>

Eglington, L. G., & Kang, S.H.K. (2018). Retrieval practice benefits deductive inference. *Educational Psychology Review*, 30(1), 215-228. <https://doi.org/10.1007/s10648-016-9386-y>

Emmerdinger, K. J., & Kuhbandner, CH. (2019). Tests improve memory – no matter if you feel good or bad while taking them. *Memory*, (27)8, 1043-1053. <https://doi.org/10.1080/09658211.2019.1618339>

Fazio, L. K (2019). Retrieval practice opportunities in middle school mathematics teachers' oral questions. *British Journal of Educational Psychology*, 89(2), 653-669. <https://doi.org/10.1111/bjep.12250>

Hostetter, A. B., Penix, E. A., Norman, M. Z., Batsell, W. R., & Carr, Th, H. (2019). The role of retrieval practice in memory and analogical problem-solving. *Quarterly Journal of Experimental Psychology* 72(4), 858–871. <https://doi.org/10.1177/1747021818771928>

Kang, S. H. K., McDermott, K. B., & Roediger, H. L. III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4-5), 528–558. <https://doi.org/10.1080/09541440601056620>

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In: Byrne, J. H. (Ed.). *Learning and Memory: A comprehensive reference*. (2nd ed., Vol. 2) (pp. 487-514). <https://doi.org/10.1016/B978-0-12-809324-5.21055-9>

Karpicke, J. D., & Aue, W., R. (2015). The testing effect is alive and well with complex materials. *Educational Psychology Review*, 27, 317-326. <https://doi.org/10.1007/s10648-015-9309-3>

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331, 772–775. <https://doi.org/10.1126/science.1199327>

Khanna, M. M. (2015). Ungraded pop quizzes: Test-enhanced learning without all the anxiety. *Teaching of Psychology*, 42, 174–178. <https://doi.org/10.1177/0098628315573144>

Larsen, D.P., Butler, A.C., & Roediger, H.L. 3rd. Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med Educ.*, 2013 Jul; 47(7), 674-82. <https://doi.org/10.1111/medu.12141>

Leahy, W., Hanham, J., & Sweller, J. (2015). High element interactivity information during problem solving may lead to failure to obtain the testing effect. *Educational Psychology Review*, 27, 265–289. <https://doi.org/10.3389/fpsyg.2018.01483>

Lyle, K. B., Bego, C. R., Hopkins, R. F., Hieb, J. L., & Ralston, P. A. (2020). How the amount and spacing of retrieval practice affect the short- and long-term retention of mathematics knowledge. *Educational Psychology Review*, 32, 277–295. <https://doi.org/10.1007/s10648-019-09489-x>

Lyle, K. B., & Crawford, N. A. (2011). Retrieving essential material at the end of lectures improves performance on statistics exams. *Teaching of Psychology*, 38(2), 94–97. <https://doi.org/10.1177/0098628311401587>

Lyle, K. B., Hopkins, R. F., Hieb, J. L., & Ralston, P. A. (2016). Spaced retrieval practice increases college students' short- and long-term retention of mathematics knowledge. *Educational Psychology Review*, 28(4), 853–873. <https://doi.org/10.1007/s10648-015-9349-8>

Niven, I., Zuckerman, H. S., & Montgomery, H. L. (1991). *An Introduction to the Theory of Numbers* (5th ed.) New York: John Wiley and Sons. Inc.

Orr, R., & Foster, S. (2017). Increasing student success using online quizzing in introductory (majors) biology. *CBE Life Sciences Education*, 12(3), 509–514. <https://doi.org/10.1187/cbe.12-10-0183>

Peterson, D., & Wissman, K. (2018). The testing effect and analogical problem-solving. *Memory*, 26(10) 1-7. <https://doi.org/10.1080/09658211.2018.1491603>

Roediger H. L., & Butler, A. C., (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1) 20-27. <https://doi.org/10.1016/j.tics.2010.09.003>

Rowland, C. A. (2014). The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect. *Psychological Bulletin*, 140(6), 1432–1463. <https://doi.org/10.1037/a0037559>

Smith, M. A., & Karpicke, J. D. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory*, 22(7), 784–802. <https://doi.org/10.1080/09658211.2013.831454>

Szöllősi, Á., Keresztes, A., Novák, B., Szászi, B., & Racsmány, M. (2017). The Testing Effect is Preserved in Stressful Final Testing Environment. *Applied Cognitive Psychology*. 31. <https://doi.org/10.1002/acp.3363>

Tran, R., Rohrer, D., & Pashler, H. (2015). Retrieval practice: The lack of transfer to deductive inferences. *Psychonomic Bulletin & Review*, 22(1), 135–140. <https://doi.org/10.3758/s13423-014-0646-x>

Tse, C.-S., & Pu, X., (2012). The effectiveness of test-enhanced learning depends on trait test anxiety and working-memory capacity. *J. Exp. Psychol. Appl.* 18 (3), 253–264. <https://doi.org/10.1037/a0029190>

van Eersel, G. G, Verkoeijen, P. P. J. L., Povilenaite, M., & Rikers, R. (2016). The testing effect and far transfer: The role of exposure to key information. *Frontiers in Psychology*, 7, Article 1977. <https://doi.org/10.3389/fpsyg.2016.01977>

van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review*, 27(2), 247–264. <https://doi.org/10.1007/s10648-015-9310-x>

Wang, Z., Lukowski, S. L, Hart, S. A., Lyons, I. M., Thompson, L. A., Kovas, Y., & Petrill, S. A. (2015). Is math anxiety always bad for math learning? The role of math motivation. *Psychological Science*, 26(12), 1863–1876. <https://doi.org/10.1177/0956797615602471>

Wong, S., S., H., Ng, G., J., P., Tempel, T. & Lim, S., W., H. (2019). Retrieval Practice Enhances Analogical Problem Solving. *The Journal of Experimental Education*, 87(1) 128–138. <https://doi.org/10.1080/00220973.2017.1409185>

Yeo, D. J., & Fazio, L. K. (2019). The optimal learning strategy depends on learning goals and processes : retrieval practice versus worked examples. *Journal of Educational Psychology*, 111(1), 73–90. <https://doi.org/10.1037/edu0000268>

Contact information: Csilla Zámbó. Eötvös Loránd University, Faculty of Primary and Pre-School Education, Department of Mathematics. 1126 Budapest, Kiss János altb. u. 40. E-mail: zambo.csilla@tok.elte.hu

Appendix A

1. Find the remainder of $2346235^{226688442}$ modulo 23.
2. Find all solutions of the following equation over the integers:

$$3x^{16} - 4y^{48} + 17z^{2013} = 34172$$

Appendix B

- 1) Determine all positive solutions of the following system of congruences below.

$$10x \equiv 5 \pmod{7} \wedge x \equiv 4 \pmod{9}$$

- 4) Prove that the equation

$$10!x^{10} + 12y^{20} + 110z^{1211} = 44z^{2017} + 6$$

has no solutions among the integers.

- 2) Find the remainder of $73737311^{99993330002}$ modulo 73.

OR

Find the remainder of $2017^{1111^{1212}}$ modulo 43.

- 5) For which positive integers n is

$$\sigma(3n) = \sigma(n) + 24$$

- 3) We know that 11 is a primitive root modulo 29. Is it true that 11^5 and 11^7 are primitive roots?

An alternative teaching and learning methodology at university to counter academic boredom

Una nueva metodología de enseñanza y aprendizaje universitarios frente al aburrimiento académico

<https://doi.org/10.4438/1988-592X-RE-2023-401-585>

Irene Sánchez-San-José

<https://orcid.org/0000-0001-7637-6424>

Universidad Complutense de Madrid

Daniel Enguídanos

<https://orcid.org/0000-0002-1110-197X>

Universidad Complutense de Madrid

Manuel Iglesias-Soilán

<https://orcid.org/0000-0003-3305-2117>

Universidad Complutense de Madrid

Patricia Ariza

<https://orcid.org/0000-0001-5829-1080>

Universidad Complutense de Madrid

Javier Aroztegui

<https://orcid.org/0000-0002-0531-1946>

Universidad Complutense de Madrid

Juan Fernández

<https://orcid.org/0000-0002-8535-5382>

Universidad Complutense de Madrid

Abstract

Boredom at university has been shown to interfere with teaching and learning. It is linked to variables such as task importance and complexity, or student's autonomy, which are in turn related to the chosen teaching and learning methodology. The goal of this study is to test the effectiveness of a new methodology based on four principles -shared teaching, flipped classroom, human model of teaching and creative evaluation- on reducing boredom. Three studies were carried out during two academic years -2019/2020 and 2020/2021-, the latter consisting of a replica aimed to test if initial data can be maintained. Method: Participants were third-year psychology students from the Complutense University of Madrid, drawn from 4 groups of the afternoon shift (49 participated in the initial study, 56 in the main study and 73 in the replica), who expressed their views using two evaluation instruments designed specifically for the purpose. Participants in the last two studies attended a total of 45 hours of the same course using the proposed methodology. Results: Results show that students overall get bored in class, with a notable reduction whenever the new methodology was used in both academic years. Conclusions: The replicated methodology fulfils its purpose. Nevertheless, more applications are convenient, enabling the study of the weight of each specific related variable in order to guarantee its correct application in other contexts.

Keywords: higher education, teaching methods, teaching, learning, boredom, active learning

Resumen

El aburrimiento, en la universidad, interfiere de forma negativa en el aprendizaje y la enseñanza. Está relacionado con variables referidas a la importancia o dificultad de la tarea o a la autonomía del estudiantado, a su vez relacionadas con la metodología utilizada. El objetivo de esta investigación fue poner a prueba hasta qué punto una nueva metodología basada en cuatro principios -docencia compartida, aula invertida, modelo humano de docente y evaluación creativa- posibilita la disminución del aburrimiento. Esta investigación consta de tres estudios realizados durante dos cursos -2019/2020 y 2020/2021-, siendo el último una réplica del principal, a fin de poner a prueba la consistencia de los resultados iniciales. Método: Los participantes fueron estudiantes de cuatro grupos del turno de tarde de 3º del Grado de Psicología de la Universidad Complutense de Madrid (49 en el estudio inicial, 56 en el principal y 73 en el de réplica) que manifestaron sus percepciones mediante dos instrumentos de valoración desarrollados ad hoc. En los dos últimos estudios, asistieron a un total de 45 horas de la asignatura en la que se empleó la metodología indicada. Resultados: Los resultados mostraron que se produce aburrimiento en las aulas y que en ambos cursos se ha conseguido una notable reducción del aburrimiento,

gracias a la nueva metodología empleada. Conclusiones: La metodología, replicada, cumple el objetivo de reducir el aburrimiento. Son, no obstante, convenientes más aplicaciones, a fin de estudiar el peso específico de cada una de las variables implicadas y para garantizar su correcta aplicación en otros contextos.

Palabras clave: Educación Superior, metodologías docentes, enseñanza, aprendizaje, aburrimiento, aprendizaje activo

Introduction

Boredom is an unpleasant and lethargic emotion characterized by a perceived lack of stimulation and often accompanied by the urge to escape the situation causing it (Pekrun et al., 2010, 2017; Van-Tilburg & Igou, 2016; Vogel-Walcutt et al., 2012). Research has shown that university classes are more likely to induce feelings of boredom compared to other situations (Chin et al., 2017; Goetz et al., 2019). A study by Mann and Robinson (2009) found that 59% of their students reported feeling bored in at least half of their classes, while 30% reported being bored in all of their classes. This sentiment is also echoed by both students and teachers in the Spanish university system, who frequently cite boredom as a problem (Iglesias-Soilán, 2020).

Despite its high prevalence, the study of academic boredom is relatively new (Sharp et al., 2020). While not all of the educational consequences of boredom are negative (Craven & Frick, 2022), this article will focus on those that are most harmful in the academic context.

Some recent meta-analyses (Camacho-Morles et al., 2021; Tze et al., 2016) show a negative relationship between boredom and academic performance. This study will only examine self-perceived performance, as evaluated by self-report, which is one of the most widely utilized methods of data collection in educational assessments (Lukas & Santiago, 2009).

This negative correlation has also been discovered between boredom and several variables that are of great significance in education, including attention, motivation, learning strategies, cognitive resources, and self-regulation (Eastwood et al., 2012; Goetz et al., 2019; Mann & Robinson, 2009; Nett et al., 2010; Sánchez-Rosas, 2019; Tze et al. 2016).

Hence, many authors have attempted to describe the antecedents of academic boredom. These seem to be more closely linked to the method in which the course is conducted, rather than other factors such as

course contents. The specialized international (Daschmann et al., 2011, 2014) and national (Fernández et al., 2021; Iglesias-Soilán, 2020) literature highlights issues like improper usage of ICT, monotony, or lack of practical relevance. Thus, the teaching and learning methodology utilized in the classroom seems to be associated with the emergence of boredom, and as a result, studying it as a key factor in reducing boredom is critical.

Currently, the lecture is the most commonly used teaching method in university education (Jiménez-Hernández et al., 2020). While its implementation can vary, it is usually characterized by the teacher taking centre stage and delivering content verbally in a one-way manner to a predominantly passive audience (Gatica-Saavedra & Rubí-González, 2021). The lecture method can reduce the perceived value of the content being taught or the activity being carried out. The subjective importance given to an activity and its potential results have been defined as one of the main antecedents of boredom, finding that low perceived value -whether accurate or not- can lead to boredom (Pekrun et al., 2017; Sharp et al., 2016).

Moreover, the absence of feedback may not allow us to know if the difficulty of the activity is appropriate for the classroom as a whole, especially in the current university context with a larger number and diversity of students (Johanns et al., 2017).

Difficulty has also been linked to the emergence of academic boredom, with findings suggesting that a significant imbalance in difficulty -the tasks being either too easy or too challenging- can lead to boredom (Acee et al., 2010; Daschmann et al., 2014; Tze et al., 2014; Westgate & Wilson, 2018).

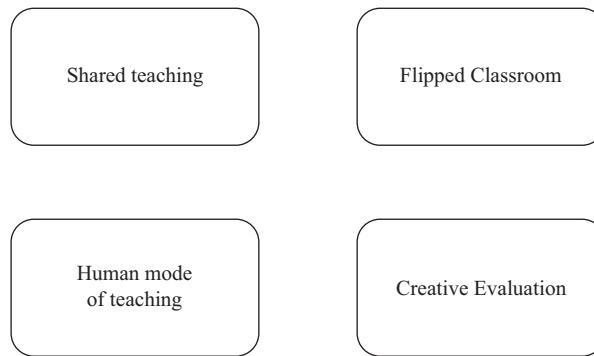
Lecture-based teaching can also reduce students' perceived control, as they are typically confined to following the teacher's instructions with limited opportunities for participation. Perceived control refers to the sense of one's ability to influence the progression and outcome of an activity. Lower perceived control has been linked to higher levels of boredom in this context (Daschmann et al., 2011; Pekrun et al., 2010; Shao et al., 2020).

To address the challenges posed by lecture-based teaching, a new teaching methodology has been developed and tested specifically for university education, with the primary aim of reducing academic boredom.

New methodology for university teaching and learning

The fundamental pillars of the new methodology are presented in Figure I: (1) shared teaching, (2) flipped classroom, (3) human mode of teaching and (4) creative evaluation (CE).

FIGURE I. The fundamental pillars of the new methodology for university teaching and learning.



Source: Compiled by author

The first pillar is shared teaching, which relies on the benefits of peer learning (Arco-Tirado et al., 2020; Stigmar, 2016) as, together with the main teacher, there is a teaching assistant - a former student of similar age to the student group who has successfully completed the subject with distinction - (Bolaños Cartujo et al., 2018; Pérez-García et al., 2020).

Although teachers possess a higher level of academic knowledge, they may sometimes experience a certain cultural and cognitive distance from their students (Lockspeiser et al., 2008), which could hinder the correct perception of the difficulty of certain tasks, potentially leading to boredom (Tze et al., 2014). Given the proximity in age of the teaching assistant and their status as a student (or recent graduate), it can be assumed that they are more familiar with the cognitive processes that students may experience - cognitive congruence - (Lockspeiser et al., 2008) and their language - shared language - (Rees et al., 2016), thus contributing to reducing the above-mentioned difficulties (Topping, 2015).

Additionally, the similarity between the teaching assistant and the students fosters a close and familiar atmosphere (Varela et al., 2015), which can enhance motivation, self-directed learning (Stigmar, 2016), and the expression of doubts in the classroom (Topping, 2015).

The second pillar of the new teaching and learning methodology is the flipped classroom approach (Bergmann & Sams, 2014). In a traditional setting, students passively receive content in the classroom (Gatica-Saavedra & Rubí-González, 2021) while practical activities are usually carried out outside the classroom (Strelan et al., 2020). The flipped classroom model shifts this approach by providing basic content before the class, often through instructional videos (Bergmann & Sams, 2014), leading to a more hands-on and active experience in the classroom (Chen et al., 2014).

In the presented methodology, students have access to lecture notes on the virtual campus and are asked to complete a specified amount of required reading before each class. This gives them greater autonomy and allows them to adapt the difficulty level to their needs, as they can decide how much time to invest in each section and whether they need to review the material, adjusting the pace to their needs. This gives students more control over their teaching and learning process, whose relationship with boredom has been noted earlier (Daschmann et al., 2011; Pekrun et al., 2010; Shao et al., 2020).

The flipped classroom approach has been shown to be more effective compared to traditional lecture-based teaching in university settings (Shi et al., 2020), and its effectiveness is further enhanced when combined with other strategies (Strelan et al., 2020).

The third pillar of the new teaching methodology is the human model, which aims to create a pleasant and close learning environment that enhances participation, focus, motivation, learning, and engagement (Varela et al., 2015) by emphasizing aspects of teacher-student communication such as empathy or accessibility (Micari & Calkins, 2021).

The fourth pillar introduces an innovative method of evaluation of the subject that actively engages all students (Rodríguez-Izquierdo, 2014), taking a more formative approach that emphasizes constant feedback for ongoing learning, rather than a purely summative perspective that focuses solely on the final outcome (Rodríguez-Gómez et al., 2012).

This form of evaluation is crucial in the methodology, due to the close relationship between the evaluation method and learning strategies (Argos et al., 2015). Small changes, such as replacing a traditional exam

with an exam that allows the use of lecture notes, can significantly shift the learning focus from memorization to understanding (Johanns et al., 2017). This is important because students who solely rely on memorization without understanding the contents may not see the relevance or value of what they have learned.

This evaluation process is referred to as “creative evaluation” (CE). In this process, each student designs their own evaluation format, which must reflect the relevant course content and the insights they have gained (e.g., reflections). This approach promotes greater adaptation to the individual and fosters autonomy in the learning process. Furthermore, students attend several tutoring sessions, where they share questions and ideas for improvement with their teachers, establishing a feedback dynamic, aligning with the principles of mediated learning as described by Feuerstein et al. (1991) and Feuerstein and Jensen (1980).

The goal of the Creative Evaluation (CE) is to enhance students' sense of control and value in the evaluation process by allowing them to select the format. This approach addresses the issue of boredom in evaluation, which has been previously noted (Daschmann et al., 2011; Pekrun et al., 2010, 2017; Shao et al., 2020; Sharp et al., 2016).

In the CE, the format itself is not assessed. Rather, the evaluation focuses on the conveyed contents, their transformation, and the inclusion of proactive elements such as summaries, examples, and reflections. The correction rubric considers criteria such as clarity, originality, coherence, quality (including attention to detail and rigor), and the significance and viability of the presented content (scientific versus subjective) (Lamont, 2009).

The main objective of this tetradimensional methodology is to reduce boredom. To achieve this goal, an initial study (study 1, boredom assessment), a main study (study 2, implementation of the methodology), and a replicating study (study 3) were conducted. Although replications are infrequent in psychology (Makel et al., 2012), some authors suggest that innovation should be supported by replication data to substantiate the consistency of initial findings (Makel & Plucker, 2014).

The three main research questions that arise from the main objective are materialized in the three studies proposed. The first question is: Does boredom really occur in the evaluated university classrooms? Secondly: Does boredom decrease after the implementation of the new methodology? And thirdly: Are the results of the main study maintained in the replication study?

Study 1: Boredom Assessment

The objective of this study was to evaluate the prevalence of boredom among students at the Faculty of Psychology of the Complutense University of Madrid. The study also explored potential solutions that students themselves could suggest to alleviate boredom.

Method

Sample

The sample for this study comprised third-year Psychology students from the afternoon group at the Complutense University of Madrid. The sample consisted of 49 students, with 40 (81.63%) females and 9 (18.37%) males. Among the total sample, 35 students were aged between 19 and 21 (71.43%), and 14 were over 21 years old (28.57%).

Instruments

An *ad-hoc* instrument was developed with the objective of gauging the opinions of students on boredom and potential solutions. It consisted of 2 items, evaluated using a Likert-type scale ranging from 1 to 7, where 1 meant “totally disagree” and 7 meant “totally agree”. The items were: “I have felt bored during my Psychology Degree classes” and “The teaching methods used in my university classes should be revised to incorporate more active and practical approaches, focusing on skill development”. The instrument also included a multiple-response item (“In the event that you have felt boredom, what would be the main causes?”) and an open-ended item where students could propose educational alternatives. The instrument comprised a total of 4 items.

Procedure

During the first week of the 2019-2020 academic year, students were given a link or QR code in the classroom that directed them to the online instrument. Descriptive analyses were carried out using SPSS version 25.0.0.1.

Results

The results of the first item suggest a propensity towards boredom ($M = 4.59$, $SD = 1.72$). The primary reasons reported for boredom in the third item were: (1) the teacher simply reads from slides during classes (76.3%), (2) the teacher does not deliver the content in an engaging manner (75.1%), (3) ineffective communication by the teacher (60.5%), and (4) a perceived lack of practical relevance (58.8%).

Regarding the second item, which concerned the necessity of adopting new teaching and learning approaches, the mean score was 6.37 ($SD = 1.07$).

Out of the 49 participants, 37 responded to the open-ended question - fourth item - highlighting the importance of more practical classes (21), alternative forms of evaluation (13), greater adaptation to diversity (8), increasing opportunities for participation (8), and the teaching of cross-functional skills (6).

Discussion

As stated in the first research question, a tendency towards boredom has been observed in university classrooms, which is consistent with findings from national and international studies (Chin et al., 2017; Goetz et al., 2019; Iglesias-Soilán, 2020). The values found are far from the supposed "ideal". Given that university studies are voluntary, boredom ratings closer to 1 rather than 4 or 5 would be expected. These levels of boredom are harmful in academic contexts, likely affecting performance, attention or motivation, among other variables (Goetz et al., 2019; Mann & Robinson, 2009; Nett et al., 2010; Sánchez-Rosas, 2019; Tze et al., 2016).

Participants indicate that the main causes of boredom are the teaching methodology and the way content is presented (such as reading from slides, ineffective communication, or unappealing presentations), as well as variables previously identified as antecedents of boredom, such as the perceived lack of practical usefulness (Pekrun et al., 2017; Sharp et al., 2016).

Given this situation, students themselves understand the need for a new methodology, as also reflected in specialized literature (Daschmann et al., 2014; Iglesias-Soilán, 2020; Mann & Robinson, 2009).

The fourth item provided information on the possible components that should make up this new methodology, which appears to be in line with what other authors have pointed out regarding boredom, advocating for a decrease in the classic format of lecture-style teaching (Gatica-Saavedra & Rubí-González, 2021) and traditional exams.

Therefore, the need for more practical activities that develop transversal competencies and facilitate the perception of the value of what is learned (Pekrun et al., 2017; Sharp et al., 2016; Stigmar, 2016), greater adaptation to the diversity of students (Johanns et al., 2017), a change in the way of evaluating, generating greater participation, and giving more control to students over their own learning process (Daschmann et al., 2011; Pekrun et al., 2010; Shao et al., 2020) becomes evident.

Study 2: Implementation of the Methodology

Method

This main study implemented the described methodology and evaluated its effectiveness.

Sample

The study included 56 participants, of whom 43 (76.79%) were female and 13 (23.21%) were male. 46 participants (82.14%) were between 19 and 22 years old, while the remaining 10 participants were over 22 years old (17.86%).

Instruments

The *ad hoc* instrument consisted of 17 items, including 13 items rated on a Likert-type scale of 1-7. These items were inspired by a Student Evaluation of Teaching (SET) model (Ching, 2018) and included statements such as “I think that creative evaluation helps me develop important transversal skills for my future.” The instrument also included a multiple-response item asking students to indicate the main causes of boredom, two dichotomous response items regarding their preference for the methodology, and a qualitative item soliciting improvement proposals.

Procedure

The study was conducted with two groups of students taking the same subject, Educational Psychology (EP). In September 2019, the methodology was presented and all doubts were resolved. The subject was then taught using the described methodology with each group until the end of January 2020, with 45 hours of instruction for each group. In the final session, the assessment instrument consisting of 17 items was administered.

Data were analysed using SPSS v.25.0.0.1, including descriptive analyses and mean differences between the initial and main study.

Results

Table I shows the main results of the first 13 items, plus the multiple response item (Figure II). Next, the results of the dichotomous items and the improvement proposals gathered from the qualitative item are presented.

TABLE I. Descriptive statistics of the 13 items with Likert-type rating scale (1-7) of the main study.

Items of the assessment instrument	Mean	Mode	S.D.
1. I have been bored in Educational Psychology (EP) class.	3.13	3	1.59
2. Generally, I have felt like going to EP class.	4.93	6	1.76
3. Generally, I felt more like going to EP than to other classes in the same course.	4.45	7	2.04
4. My interest in EP is due to the methodology.	5.00	5	1.83
5. I have found the EP class model to be more useful than other models I have experienced during this term.	5.41	7	1.68
6. The EP classes have been practical.	5.46	5	1.35
7. I have acquired transversal competencies during EP.	5.90	7	1.41
8. I believe that Creative Evaluation (CE) provides me with more benefits than other forms of evaluation.	5.55	6	1.65
9. I believe that CE helps me develop other skills important for my future.	5.73	7	1.61
10. I believe that what I have learned through CE will be better retained over time than what I have learned through assessment by exam.	5.79	7	1.46
11. I think that CE is easier (requiring less effort and work) than other types of assessment.	3.14	2	1.86
12. I believe that a human model of teaching favours my learning.	6.50	7	.69

(Continued)

TABLE I. Descriptive statistics of the 13 items with Likert-type rating scale (1-7) of the main study (Continued)

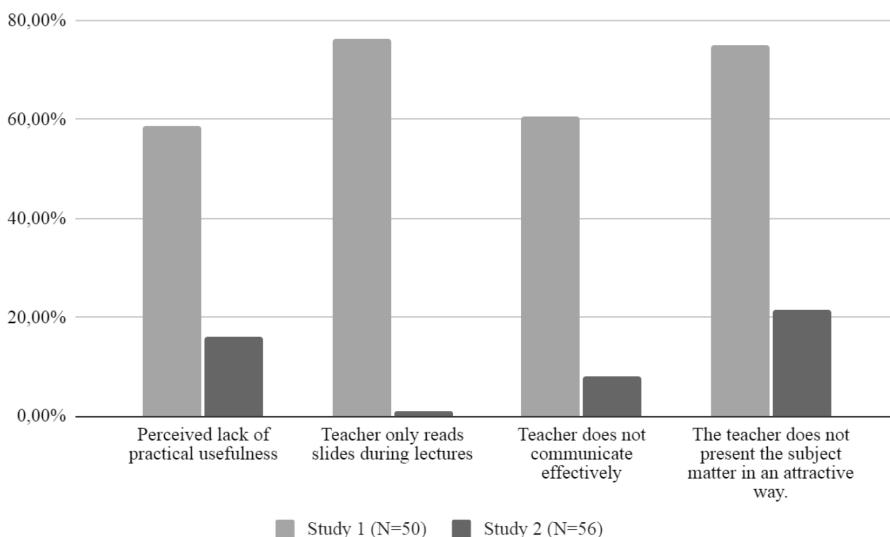
Items of the assessment instrument	Mean	Mode	S.D.
13. I believe that the EP teacher has established a human model (empathy, closeness, respect, understanding...) as opposed to a purely cognitive model (learning concepts, advancement in the subject...).	6.25	7	.98

Source: Compiled by author

First, it is noteworthy that participants showed much higher levels of boredom ($p < .000$, $d = .816$) throughout their studies ($M_{Study1} = 4.59$, $S.D._{Study1} = 1.72$) compared to levels after the application of the new methodology (item 1; $M_{Study2} = 3.13$, $S.D._{Study2} = 1.59$).

Changes in boredom rating levels were also found between baseline and main study (Figure II).

FIGURE II. Opinions on the causes of academic boredom throughout the Psychology Degree versus the new methodology.



Source: Compiled by author

As for the rest of the quantitative items (2-13), there is a high interest in attending the course (items 2 and 3), which is explicitly related to the methodology used (item 4).

Participants perceived the classes as useful (item 5), this being possibly related to their practical nature (item 6), which seemed to foster the development of competencies (item 7).

Regarding the assessment model of the subject, students indicated that CE provided them with more benefits than other types of evaluation (item 8) and that it favoured their learning in different aspects (items 9 and 10), even assuming that CE was not easier than other types of assessment or that it required less effort (item 11).

In addition, students indicated that the relationship model with the faculty was a human model that favoured their learning (items 12 and 13).

Regarding the two dichotomous response items, 91.07% of the students indicated that, if they could choose the evaluation format again, they would opt for CE over a traditional exam. The same value was given for choosing the new methodology again, indicating high satisfaction.

The qualitative item made it possible to collect proposals for improving the methodology. Some students pointed out, as relative problems, the uncertainty caused by the high flexibility (6) and the lack of time to delve into some aspects (5).

Discussion

The second research question examined whether the new methodology reduced boredom. The findings indicate a significant decrease in boredom after the implementation of the new methodology. Students attribute their increased interest in the subject to the new methodology, which is consistent with both the initial study and existing literature. The results also show a noticeable reduction in the boredom levels previously associated with lectures (Daschmann et al., 2014; Fernández et al., 2021; Iglesias-Soilán, 2020; Mann & Robinson, 2009).

The participants expressed a positive attitude towards the practical nature of the new methodology that enables them to develop transversal competencies. They also appreciated the CE approach, which provides greater adaptability to diverse student needs and enables more control over their learning and evaluation process. These findings are consistent with the feedback obtained from students in the first study and are supported by relevant literature (Acee et al., 2010; Daschmann et al., 2011; Pekrun et al., 2010, 2017; Shao et al., 2020; Sharp et al., 2016; Westgate & Wilson, 2018).

On the other hand, students positively value the human model that they reported having found in the course. This seems to have favoured their learning, as is also reflected in the literature (Varela et al., 2015).

The analysis of the open-ended question revealed areas for improvement, including the uncertainty arising from the flexibility of the new methodology and a perceived lack of time to fully engage with some course content.

The issue of flexibility warrants further investigation, as it is unclear whether the perceived flexibility of the new methodology has been excessive, having a potentially negative impact, or if students are simply not accustomed to this type of learning approach, having experienced more directive academic contexts in the past. Similarly, the identified lack of time could be due to either an insufficient amount of time allocated for course content or an increased interest in the subject matter, prompting students to dedicate more time to the course. These student feedback points require further exploration in future studies.

Study 3: Replication

The primary objective of this study was to replicate the findings of Study 2 and evaluate the consistency of the results.

Method

Sample

There were 73 participants in the replication: 58 females (79.5%), 14 males (19.2%) and 1 person who did not indicate their sex. 64 were between 19 and 22 years old (87.7%) and 9 were over 22 years old (12.3%).

Instruments

The study employed the same instrument as Study 2, with the exclusion of the qualitative item, consisting of 16 items.

Procedure

The study followed the same procedures as Study 2. Notably, during the academic year in which this study was conducted, classes were delivered on a blended basis, with alternating in-person and online sessions every other week. This arrangement was implemented due to the COVID-19 health measures.

SPSS v.25.0.0.1 was used for data analysis. Descriptive analyses were performed. In addition, the items of Study 2 and 3 were contrasted, using a two-way comparison of means: statistically significant differences and the TOST (Two One-sided Tests) of statistically significant equivalence (Lakens et al., 2018). The Two One-Sided Test (TOST) approach is used to determine an equivalence between two groups by conducting two separate hypothesis tests. The null hypothesis for the first test is that the second measure is lower than the first one, while the null hypothesis for the second test is that the second measure is greater than the first one. Equivalence is established only when both null hypotheses can be rejected with a probability of erroneous rejection of less than $p < .05$, indicating a statistically significant equivalence with certain interpretive nuances.

In research, it is common to use statistical tests to identify significant differences, particularly in cases where the goal is to evaluate the effectiveness of an intervention or treatment. A statistically significant result ($p < .05$) suggests that the observed difference is unlikely to have occurred by chance. On the other hand, when there is no statistically significant difference, it is challenging to draw conclusions because the data may be attributable to chance.

In this replication study, one possible outcome is that the data obtained are equivalent to or even better than those in the original study. If the data are equivalent, we would need a statistical test to determine whether the similar values are unlikely ($p < .05$) to have occurred by chance, known as the test for statistically significant equivalence.

If the results are either maintained or improved in the replication study, it will be considered a favourable outcome. Improvement in this case refers to the existence of statistically significant differences that indicate better results than those found in the original study.

Results

To verify that the conditions were similar to those of the main study, boredom was measured again at the beginning of the course. The replicate participants also indicated high boredom.

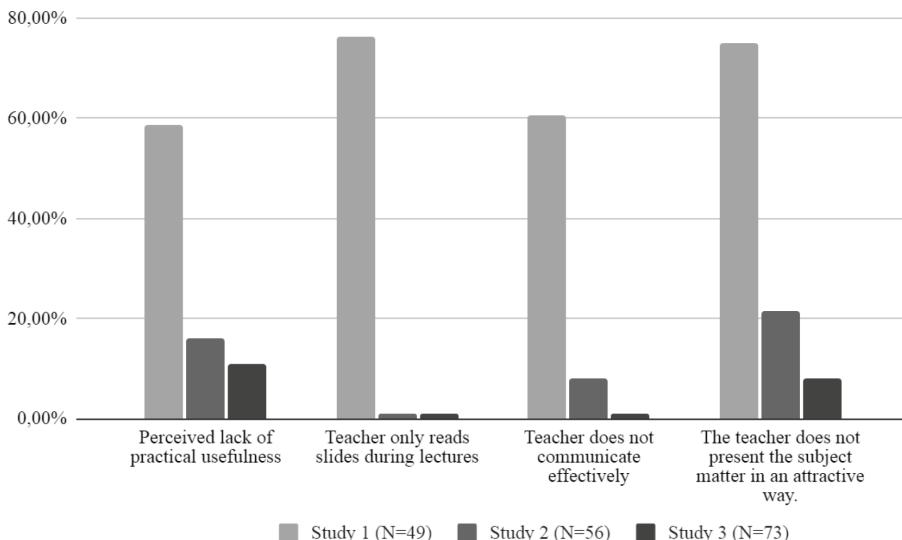
Table II shows the main results of the 13 items with Likert-type rating scales. Next, the multiple-response item referring to the different ratings of the causes of boredom is shown, comparing the three studies (Figure III). Finally, comparisons between the main study and the replication are presented.

TABLE II. Descriptive statistics of the 13 items with Likert-type rating scale (1-7) of the replication.

Assessment instrument items (short form)	Mean	Mode	S.D.
1. I have been bored in EP.	2.56	2	1.20
2. I felt like going to EP.	5.40	7	1.39
3. I have felt more like going to EP than to other classes.	5.26	7	1.58
4. My interest is due to the methodology.	5.47	7	1.45
5. This model is more useful than others.	6.03	7	1.01
6. The classes are practical.	5.55	5	1.03
7. I have acquired competencies.	5.56	6	1.37
8. CE provides more benefits than other evaluations.	6.04	7	1.09
9. I have acquired competencies through CE.	6.03	7	1.07
10. CE promotes learning.	6.12	7	1.28
11. CE is easier than other assessments.	2.27	2	1.57
12. The human model favours my learning.	6.56	7	.73
13. I perceived a human model in EP.	6.41	7	.76

Source: Compiled by author.

FIGURE III. Opinions on the causes of classroom boredom in the three studies.

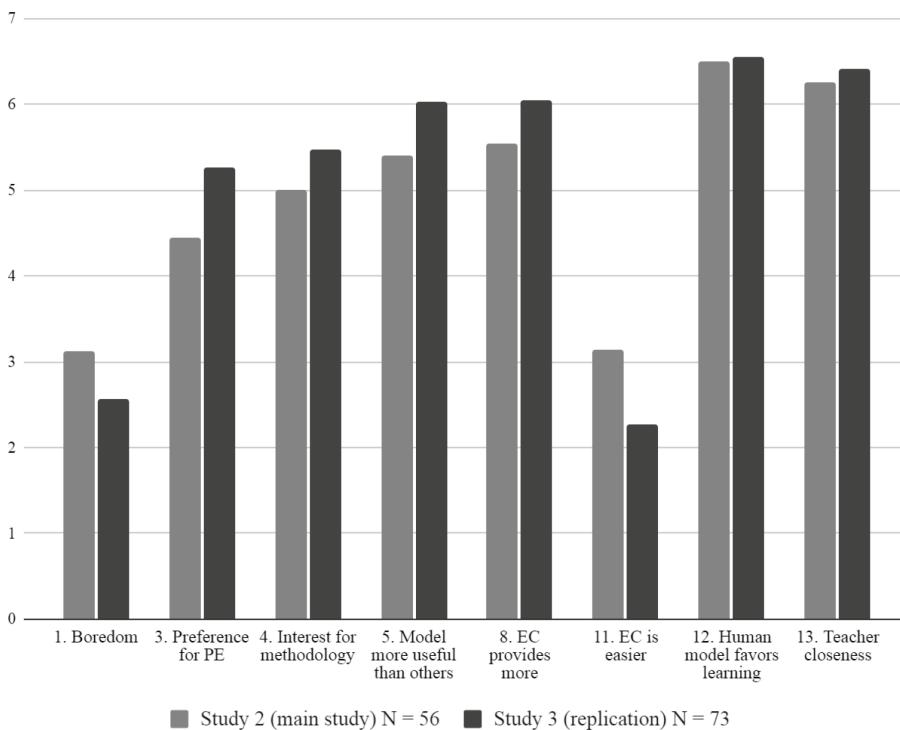


Source: Compiled by author.

The replication study shows a significant decrease in boredom, which may be explained by the evaluation of the perceived causes of boredom compared to the previous studies. The high interest in attending the course (items 2 and 3) is related to the methodology (item 4), which is perceived as useful and practical and promotes the development of transversal competencies (items 5, 6 and 7). Creative Evaluation continues to stand out as a model that enhances learning, despite requiring more work and effort than traditional systems (items 8, 9, 10 and 11). The human relationship model also continues to positively affect learning (items 12 and 13). These findings are consistent with the results of study 2, as shown in Figure IV.

The comparative results of the results of studies 2 and 3, with respect to the positive impact of the methodology, are shown in Table III. The items with favourable -similar or better- results are those that show: (1) statistically significant equivalence (e.g., item 1); (2) statistically significant differences with better results in the replication study (e.g., item 5) or (3) lie in an intermediate situation (e.g., item 2). A result will be considered to be better when the value obtained in the replication is higher, except for items 1 (boredom) and 11 (ease of CE), where a lower value will be better.

FIGURE IV. Descriptive comparison of the means of some of the most relevant items in studies 2 and 3.



¹Items have been summarized.

²Source: Compiled by author.

As can be seen, there is high consistency between the values obtained in studies 2 and 3, with trends of improvement in the replication with respect to the main study. Thus, the vast majority of the results support the maintenance or improvement of the effects in the replication, with the exception of item 7 (transversal competencies).

Regarding the usefulness of the model compared to others (item 5), the results are better in the replication. Regarding CE, participants in the replication perceive it as more difficult than those in the main study, although 84.90% of the participants would choose CE again over a traditional exam (91.07% in study 2).

In addition, 95.9% would again select this methodology over a traditional one (91.07% in study 2).

TABLE III. Favourable results (equivalent or better) from the replication study of the application of the new methodology.

ABBREVIATED ITEMS	STATISTICALLY SIGNIFICANT EQUIVALENCE		IMPROVE-MENT WITHOUT REACHING STATISTICAL SIG.	STATISTICALLY SIGNIFICANT FAVOURABLE DIFFERENCES	
	t	P		$M_{Study\ 3}$ better	T
1. I have been bored in EP.	$t>(107.96)= 4.01$ $t<(107.96)= -1.51$	(.000, .007)*	Yes	$t(107.96)= 1.25$.214
2. I felt like going to EP.	$t>(102.56)= 1.09$ $t<(102.56)= -4.37$	(.139, .000)	Yes ¹	$t(102.56)= -1.64$.104
3. I have felt more like going to EP than to other classes.	$t>(101.16)= 0.25$ $t<(101.16)= -5.20$	(.402, .000)	Yes ¹	$t(101.16)= -2.47$.121
4. My interest is due to the methodology.	$t>(102.78)= 1.17$ $t<(102.78)= -4.29$	(.123, .000)	Yes ¹	$t(102.78)= -1.56$.121
5. This model is more useful than others.	$t>(84.88)= 0.22$ $t<(84.88)= -5.07$	(.414, .000)	Yes	$t(84.88)= -2.43$.017**
6. The classes are practical.	$t>(99.84)= 2.33$ $t<(99.84)= -3.10$	(.101, .001)	Yes ¹	$t(99.84)= -0.39$.700
7. I have acquired competencies.	$t>(116.93)= 4.14$ $t<(116.93)= -1.47$	(.000, .072)	No	$t(116.93)= 1.34$.184
8. The CE provides more benefits than other evaluations.	$t>(90.00)= 0.75$ $t<(90.00)= -4.59$	(.226, .000)	Yes ¹	$t(90.00)= -1.91$.059
9. I have acquired competencies through CE.	$t>(90.36)= 1.49$ $t<(90.36)= -3.86$	(.070, .000)	Yes ¹	$t(90.36)= -1.19$.239
10. CE promotes learning.	$t>(109.73)= 1.39$ $t<(109.73)= -4.14$	(.083, .000)	Yes ¹	$t(109.74)= -1.37$.173
11. CE is easier than other assessments.	$t>(106.81)= 5.56$ $t<(109.73)= 0.06$	(.000, .059)	Yes	$t(106.81)= 2.81$.006**
12. The human model favours my learning.	$t>(121.47)= 2.34$ $t<(121.47)= -3.33$	(.010, .001)*	Yes	$t(121.47)= -0.49$.623
13. I perceived a human model in EP.	$t>(101.32)= 1.71$ $t<(101.32)= -3.74$	(.046, .000)*	Yes	$t(101.32)= -1.02$.311

^{1*} Statistically significant equivalence with $p < .05$. ¹ Better result in the replication study, without reaching a statistically significant difference. ^{2**} Statistically significant differences with $p < .05$ and better result in the replication.

³ Source: Compiled by author

Discussion

In relation to the third research question posed - are the results of previous studies maintained? -it is observed that the results initially obtained are maintained in this study or even improved in some cases. The maintenance of the results in a context of blended teaching reflects the adaptability of the methodology to other contexts, maintaining its effectiveness.

Assessments on the causes of boredom and the relevance of the new methodology as an effective means to reduce it are highly consistent, which is in line with findings from previous research (Chin et al., 2017; Daschmann et al., 2014; Goetz et al., 2019; Iglesias-Soilán, 2020; Mann & Robinson, 2009). These results highlight the importance of proposing alternative methodologies to traditional ones (Jiménez-Hernández et al., 2020).

The reasons for the improved results in the replication study compared to the main study are uncertain. Several possibilities exist, such as (1) a comparative effect with other subjects that predominantly rely on lectures and have been more negatively affected by blended teaching, or (2) an enhancement in the application of the methodology. Further research is needed to explore these potential factors.

Conclusions

The first objective of this study was to investigate the prevalence of boredom in university classrooms, as previous research has highlighted its negative impact on the academic environment (Camacho-Morles et al., 2021; Eastwood et al., 2012; Goetz et al., 2019; Nett et al., 2010; Sánchez-Rosas, 2019; Tze et al. 2016). The findings from the three studies support the existing literature, which shows that boredom is prevalent among the university population (Mann & Robinson, 2009; Iglesias-Soilán, 2020) and that its levels are far from the ideal for effective learning.

According to both the students themselves and the national and international literature, boredom in university classrooms seems to be caused mainly by factors such as a perceived lack of usefulness in the subjects, a perceived lack of control over learning, or difficulty adapting to diversity (Acee et al., 2010; Daschmann et al., 2011, 2014; Pekrun et al., 2010,

2017; Shao et al., 2020; Sharp et al., 2016; Tze et al., 2014; Westgate & Wilson, 2018). These findings indicate a need for a change in teaching and learning methodology to address the issue of boredom in the academic environment.

An alternative methodology was proposed and it has shown high effectiveness in considerably reducing high levels of academic boredom. This methodology is based on four pillars, the individual influence of which remains unknown and requires further research. However, these pillars are believed to have contributed to the success of the methodology.

To ensure the reliability of the initial findings, a replication study was conducted. Replication is a critical component of the scientific process that helps to verify the validity of research results. However, in the social sciences, replication studies are not as common as they should be (Makel et al., 2012; Makel & Plucker, 2014). This is partly due to the *ad novitatem* fallacy, which involves perceiving new research as superior solely because of its novelty, often resulting in its preferential treatment during the peer-review process.

In this case, the replication study has consolidated the results of the main study, showing similarities and even improvements. These findings suggest the need to continue applying and refining the methodology to increase its effectiveness while reducing the required time and resources.

It is important to note that a more rigorous replication would involve professionals other than those who conducted the initial research. Therefore, it is advisable to continue with replications in diverse groups, locations, and subjects, inviting other professionals to evaluate the effectiveness of the methodology. These replications should include additional measures of variables such as boredom and performance, building upon the initial student perception as a first step in the study of these variables. Furthermore, the integration of qualitative data collection techniques would enable exploration of other antecedents and consequences of boredom.

In summary, the challenge at hand is to try to respond to the need to create alternatives for university teaching and learning processes that can better satisfy both teachers and students and reduce the harmful levels of boredom that currently exist in university classrooms. Some of these aspects are already being investigated in a doctoral thesis.

Acknowledgment

To all the newly joined faculty-researcher team, especially Sandra San José Pérez, whose attention to detail has become indispensable to us.

Bibliographical references

Acee, T. W., Kim, H., Kim, H. J., Kim, J. I., Chu, H. N. R., Kim, M., Cho, Y., Wicker, F. W., & Boredom Research Group. (2010). Academic boredom in under-and over-challenging situations. *Contemporary Educational Psychology, 35*(1), 17-27. <https://doi.org/csm7zm>

Arco-Tirado, J. L., Fernández-Martín, F. D., & Hervás-Torres, M. (2020). Evidence-based peer-tutoring program to improve students' performance at the university. *Studies in Higher Education, 45*(11), 2190-2202. <https://doi.org/gfzc9z>

Argos, J., Ezquerro, P., Osoro, J. M., Salvador, L., & Castro, A. (2015). La evaluación de los aprendizajes de los estudiantes en el marco del Espacio Europeo de Educación Superior (EEES): sus prácticas, preferencias y evolución. *European Journal of Investigation in Health, Psychology and Education, 3*(3), 181-194. <https://doi.org/hk57>

Bergmann, J., & Sams, A. (2014). *Flipped learning: Gateway to student engagement*. International Society for Technology in Education.

Bolaños Cartujo, I., El-Seyoufi Cortés, S., Marqués Delgado, F., Casado Martínez, M. D. P., Escorial Martín, S., & Gómez Ruiz, M. S. (2018). *Proyecto de auxiliares docentes pregraduados*. Proyecto Nº 38 INNO-DOCENCIA 2017 UCM. <https://cutt.ly/4StnvoU>

Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review, 33*(3), 1051-1095. <https://doi.org/gntx4p>

Chen, Y., Wang, Y., & Chen, N. S. (2014). Is FLIP enough? Or should we use the FLIPPED model instead? *Computers & Education, 79*, 16-27. <https://doi.org/gf4pdg>

Chin, A., Markey, A., Bhargava, S., Kassam, K. S., & Loewenstein, G. (2017). Bored in the USA: Experience sampling and boredom in everyday life. *Emotion, 17*(2), 359-368. <https://doi.org/f9zrjt>

Ching, G. (2018). A literature review on the student evaluation of teaching: An examination of the search, experience, and credence qualities of SET. *Higher Education Evaluation and Development*, 12(2), 63-84. <https://doi.org/hk59>

Craven, A., & Frick, L. (2022). Boredom as a basis for fostering creativity in higher education: A call for pedagogical bravery. *Innovations in Education and Teaching International*, 1-13. <https://doi.org/jm7b>

Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2011). Testing the predictors of boredom at school: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology*, 81(3), 421-440. <https://doi.org/dnvf67>

Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2014). Exploring the antecedents of boredom: Do teachers know why students are bored? *Teaching and Teacher Education*, 39, 22-30.

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on Psychological Science*, 7(5), 482-495. <https://doi.org/f399ds>

Fernández, J., Ariza, P., Enguídanos, D., Iglesias-Soilán, M., Aroztegui, J., & Sánchez-San-José, I. (2021). *Nueva metodología docente dinámica frente al aburrimiento en las clases: una réplica*. Proyecto N°338 INNOVA-DOCENCIA 2020 UCM. <https://cutt.ly/TStmdPu>

Feuerstein, R., & Jensen, M. R. (1980). Instrumental enrichment: Theoretical basis, goals, and instruments. *The Educational Forum*, 44(4), 401-423.

Feuerstein, R., Klein, P. S., & Tannenbaum, A. J. (1991). *Mediated Learning Experience (MLE): Theoretical, psychosocial and learning implications*. Freund.

Gatica-Saavedra, M., & Rubí-González, P. (2021). La clase magistral en el contexto del modelo educativo basado en competencias. *Revista Electrónica Educare*, 25(1), 1-12. <https://doi.org/hk6d>

Goetz, T., Hall, N. C., & Krannich, M. (2019). Boredom. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge handbook of motivation and learning* (pp. 465-489). Cambridge University Press.

Iglesias-Soilán, M. (2020). La enseñanza universitaria: el aburrimiento en las aulas. *PUBLICACIONES*, 50(3), 93-124. <https://doi.org/hk6f>

Jiménez-Hernández, D., González-Ortiz, J. J., & Tornel-Abellán, M. (2020). Metodologías activas en la universidad y su relación con los enfoques

de enseñanza. *Profesorado, Revista de Currículum y Formación del Profesorado*, 24(1), 76-94. <https://doi.org/hk6g>

Johanns, B., Dinkens, A., & Moore, J. (2017). A systematic review comparing open-book and closed-book examinations: Evaluating effects on development of critical thinking skills. *Nurse Education in Practice*, 27, 89-94. <https://doi.org/gcm7kg>

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/gdj7s9>

Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment*. Harvard University Press.

Lockspeiser, T. M., O'Sullivan, P., Teherani, A., & Muller, J. (2008). Understanding the experience of being taught by peers: The value of social and cognitive congruence. *Advances in Health Sciences Education*, 13(3), 361-372. <https://doi.org/fjf746>

Lukas, J. F., & Santiago, K. (2009). *Evaluación educativa* (2^a ed.). Alianza.

Makel, M. C., & Plucker, J. A. (2014). Creativity is more than novelty: Reconsidering replication as a creativity act. *Educational Researcher*, 43(6), 304-316. <https://doi.org/gmm29j>

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7(6), 537-542. <https://doi.org/f4fdbt>

Mann, S., & Robinson, A. (2009). Boredom in the lecture theatre: An investigation into the contributors, moderators and outcomes of boredom amongst university students. *British Educational Research Journal*, 35(2), 243-258. <https://doi.org/b67dcr>

Micari, M., & Calkins, S. (2021). Is it OK to ask? The impact of instructor openness to questions on student help-seeking and academic outcomes. *Active Learning in Higher Education*, 22(2), 143-157.

Nett, U. E., Goetz, T., & Daniels, L. M. (2010). What to do when feeling bored?: Students' strategies for coping with boredom. *Learning and Individual Differences*, 20(6), 626-638.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3), 531-548. <https://doi.org/fqjktm>

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal

models of reciprocal effects. *Child Development*, 88(5), 1653-1670. <https://doi.org/gbw6vg>

Pérez García, E.M., Bolaños Cartujo, J.I., Barrón López De Roda, A., Casado Martínez, M. P., Iglesias-Soilán, M., Martín Dobón, E., Enguídanos, D., Escorial Martín, S., Fernández González, S., Gómez Ruiz, M.S., Ondé Pérez, D., Orio Ortiz, L., & Río Grande, D. P., (2020). *Programas de auxiliares docentes posgraduados*. Proyecto Nº 269 INNOVACIÓN DOCENCIA 2020 UCM. <https://cutt.ly/HStYrXw>

Rees, E. L., Quinn, P. J., Davies, B., & Fotheringham, V. (2016). How does peer teaching compare to faculty teaching? A systematic review and meta-analysis. *Medical Teacher*, 38(8), 829-837. <https://doi.org/gkscf7>

Rodríguez-Gómez, G., Ibarra-Saiz, M., Gallego-Noche, B., Gómez-Ruiz, M., & Quesada-Serra, V. (2012). La voz del estudiante en la evaluación del aprendizaje: un camino por recorrer en la universidad. *RELIEVE, Revista Electrónica de Investigación y Evaluación Educativa*, 18(2), 1-21. <https://doi.org/c897>

Rodríguez-Izquierdo, R. M. (2014). Modelo formativo en el Espacio Europeo de Educación Superior: valoraciones de los estudiantes. *Aula Abierta*, 42(2), 106-113. <https://doi.org/f2sr92>

Sánchez-Rosas, J. (2019). Aburrimiento: función, causas y estrategias de intervención. In S. Vernucci & E. Zamora (Eds.), *La ciencia de enseñar. Aportes desde la psicología cognitiva a la educación* (pp. 21-30). Universidad Nacional de Mar del Plata.

Shao, K., Pekrun, R., Marsh, H. W., & Loderer, K. (2020). Control-value appraisals, achievement emotions, and foreign language performance: A latent interaction analysis. *Learning and Instruction*, 69, 101356. <https://doi.org/fmpq>

Sharp, J. G., Hemmings, B., Kay, R., Murphy, B., & Elliott, S. (2016). Academic boredom among students in higher education: A mixed-methods exploration of characteristics, contributors and consequences. *Journal of Further and Higher Education*, 41(5), 657-677. <https://doi.org/gktjdd>

Sharp, J. G., Sharp, J. C., & Young, E. (2020). Academic boredom, engagement and the achievement of undergraduate students at university: A review and synthesis of relevant literature. *Research Papers in Education*, 35(2), 144-184. <https://doi.org/ggpt8v>

Shi, Y., Ma, Y., MacLeod, J., & Yang, H. H. (2020). College students' cognitive learning outcomes in flipped classroom instruction: A meta-analysis

of the empirical literature. *Journal of Computers in Education*, 7(1), 79-103. <https://doi.org/fbbf>

Stigmar, M. (2016). Peer-to-peer teaching in higher education: A critical literature review. *Mentoring & Tutoring: Partnership in Learning*, 24(2), 124-136. <https://doi.org/f3r389>

Strelan, P., Osborn, A., & Palmer, E. (2020). Student satisfaction with courses and instructors in a flipped classroom: A meta-analysis. *Journal of Computer Assisted Learning*, 36(3), 295-314. <https://doi.org/hk6p>

Topping, K. (2015). Peer tutoring: Old method, new developments. *Journal for the Study of Education and Development*, 38(1), 1-29. <https://doi.org/hk6q>

Tze, V.M.C., Daniels, L.M., & Klassen, R.M (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review*, 28(1), 119-144. <https://doi.org/gdvp59>

Tze, V. M., Klassen, R. M., & Daniels, L. M. (2014). Patterns of boredom and its relationship with perceived autonomy support and engagement. *Contemporary Educational Psychology*, 39(3), 175-187. <https://doi.org/gft3dh>

Van-Tilburg, W. A., & Igou, E. R. (2016). Boredom begs to differ: Differentiation from other negative emotions. *Emotion*, 17(2), 309-332. <https://doi.org/f9w6fj>

Varela, O. E., Cater III, J. J., & Michel, N. (2015). Learner-instructor similarity: A social attribution approach to learning. *Journal of Management Development*, 34(4), 460-475. <https://doi.org/hk6r>

Vogel-Walcutt, J. J., Fiorella, L., Carper, T., & Schatz, S. (2012). The definition, assessment, and mitigation of state boredom within educational settings: A comprehensive review. *Educational Psychology Review*, 24(1), 89-111. <https://doi.org/d5tfqk>

Westgate, E. C., & Wilson, T. D. (2018). Boring thoughts and bored minds: The MAC model of boredom and cognitive engagement. *Psychological Review*, 125(5), 689-713. <https://doi.org/gffhgc>

Contact address: Irene Sánchez-San-José. Universidad Complutense de Madrid, Facultad de Psicología, Sección Departamental de Investigación y Psicología en Educación. Campus de Somosaguas, Ctra. de Húmera, s/n, 28223 Pozuelo de Alarcón, Madrid. E-mail: isanch16@ucm.es

Influence of Physical Education on physical activity levels in adolescence. A systematic review

Influencia de la Educación Física en los niveles de actividad física en la adolescencia. Una revisión sistemática

<https://doi.org/10.4438/1988-592X-RE-2023-401-586>

Ariadna Hernaiz-Sánchez

<https://orcid.org/0000-0002-6718-2887>

Universidad Europea de Madrid

Andrea Bäder-Gilabert

<https://orcid.org/0000-0002-4289-3631>

Universidad Europea de Madrid

Abstract

Lack of physical activity, especially in teenagers, has become a major problem worldwide. More than 80% of the adolescent population does not meet the minimum recommendations set by the WHO. To combat low levels of physical activity, Physical Education could be one tool. However, it does not seem to be contributing to improving these percentages. The main objective of this systematic review was to look into the scientific literature so as to analyse the influence of PE sessions on the physical activity levels of adolescents. It was aimed at identifying whether the average levels of moderate or vigorous physical activity in physical education sessions reached the minimum recommended levels, investigating the factors that interfere with the results of these levels and finding out whether the effects of the Spanish secondary physical education curriculum on the physical activity levels of adolescents are within the standards. Twenty-one articles extracted from three different databases were analysed using the same search equation. Physical education has not achieved the minimum percentages established in secondary school; the main variables that have interfered in these results were gender and the contents of the lessons. The results of the studies carried out in Spain have been within the medium threshold. Given the

importance of physical activity levels on the health of any individual, further research is needed. Thus, it would be appropriate for future studies to be based on the same method to avoid the differences that have been found.

Keywords: physical activity, physical education, moderate or vigorous physical activity (MVPA), adolescence, physical activity level.

Resumen

La falta de actividad física, sobre todo en las poblaciones adolescentes se ha convertido en un problema de gran envergadura a nivel mundial. Más de un 80% de la población adolescente no cumple con las recomendaciones mínimas establecidas por la OMS. Para tratar de combatir los escasos niveles de actividad física, la Educación Física (EF) podría ser una herramienta. Sin embargo, no parece estar contribuyendo a mejorar estos porcentajes. El objetivo principal de esta revisión sistemática ha sido revisar la literatura científica para analizar la influencia que ejercen las sesiones de EF en los niveles de actividad física de los adolescentes. Se ha pretendido identificar si los niveles medios de actividad física moderada o vigorosa en las sesiones de EF alcanzaba los mínimos recomendados, investigar los factores que interfieren en los resultados de estos niveles y conocer si los efectos del currículum de la EF de secundaria de España sobre los niveles de actividad física de los adolescentes están dentro de los estándares. Se han analizado 21 artículos extraídos de tres bases de datos distintas a través de la misma ecuación de búsqueda. El estudio ha concluido que la EF no ha logrado los porcentajes mínimos establecidos en secundaria, las principales variables que han interferido en estos resultados han sido el género de los adolescentes y los contenidos de las lecciones. Los resultados de los estudios realizados en España se han situado dentro del umbral medio. Dada la importancia de los niveles de actividad física sobre la salud de cualquier individuo, será necesario seguir indagando en la temática. Sería adecuado que los futuros estudios se basen en un mismo método para evitar las diferencias que se han encontrado.

Palabras clave: actividad física, educación física, actividad física moderada o vigorosa (MVPA), adolescencia, nivel de actividad física.

Introduction

Several studies show evidence of the negative effects of a lack of physical activity on the health of any individual (Guthold et al., 2018). Regular physical activity helps in the prevention of different chronic, and cardiovascular conditions, as well as diabetes, cancer, mental health and other

diseases (Bauman, 2004; Biddle & Asare, 2011). For this reason, the subject of Physical Education (PE) has an added responsibility that will not only affect an individual's adolescence but will be reflected throughout the course of his or her adult life (Haerens et al., 2010). The practice of physical activity is directly and positively related to the health of any individual (Martins et al., 2018).

In order to raise awareness of the importance of being active during the different stages of life, the World Health Organization (WHO) established a clear and summarised set of recommendations. These were adapted to different age groups considering the metabolic needs of the time. For young people aged 5 to 17, it established a minimum of 60 minutes of moderate or vigorous physical activity per day, and for those aged 18 to 64 a minimum of 150 minutes of moderate aerobic physical activity or 75 minutes of vigorous aerobic physical activity per week (World Health Organization, 2010).

Currently, 81% of adolescents and 27.5% of adults do not meet these guidelines (Guthold et al., 2018, 2020). In the daily life of teenagers, there are different actions that directly and indirectly affect these parameters, but it is necessary to highlight that most of their time is spent inside schools or high schools. For this reason, the educational environment becomes the ideal setting to close the gap with the WHO. This is where PE takes on a responsibility for these values (Murillo et al., 2014). Specifically, as cited in Royal Decree 1105/2014, of 26th December, which establishes the basic curriculum for Compulsory Secondary Education and the Baccalaureate, PE professionals aim to consolidate the regularity of physical activity and avoid a sedentary lifestyle.

Various organisations advise that, during PE sessions in secondary school, a minimum of 50% of the time should be devoted to moderate or vigorous physical activity (Association for Physical Education, 2015) to ensure a minimum level of physical activity during those days students take the subject. However, several studies have shown that these requirements are not met during the sessions (Cantero et al., 2019; Fairclough & Stratton, 2005; Mayorga-Vega et al., 2018; Molina-García et al., 2016; Murillo et al., 2014; Dudley et al. 2012). Yet, there is general agreement that all adolescents, regardless of their singularities, reduce the distances with the marked suggestions on days when they have PE class (Groffik et al., 2020; Mayorga-Vega et al., 2018). In this way, this subject remains one of the main channels to help them reach the established minima. At the same time, it is of vital importance that during the practice of

physical activity they extract positive feelings, since, throughout adolescence, activity patterns are established that extend well into adulthood (Whitehead, 2010; Martins et al., 2018; Zhou & Wang, 2019).

In order to try to reverse the trend towards sedentary lifestyles, further research is needed on the different variables involved in the influence of PE on the physical activity level of adolescents with the aim of closing the gap with the indicated parameters. These variables may derive from different factors that can be classified as: inter-individual, those specific to the subject; pedagogical, those related to the educational environment; or institutional, those triggered by governmental decisions (Delextrat et al., 2020).

Inter-individual factors

One of the most common and significant variables among these in the different articles is gender. It has been found that boys reach higher amounts of moderate or vigorous physical activity than girls (Cantero et al., 2019; Fairclough & Stratton, 2005; Groffik et al., 2020; Mayorga-Vega et al., 2018; Molina-García et al., 2016; Murillo et al., 2014; Zhou & Wang, 2019). There are several factors responsible for these results, but the most commonly mentioned was the type of exercises performed in the PE sessions. These directly affect the level of physical activity, not only because of their demand, but also because of the motivation they trigger in adolescents when performing them. Generally, girls prefer activities in which they can feel included and which they can do cooperatively at a lower intensity. In contrast, most boys prioritise competitive situations at a higher intensity (Fairclough & Stratton, 2005; Martins et al., 2018).

Reference should also be made to the students' physical fitness, with those with a higher percentage of body fat performing less physical activity compared to adolescents with favourable anthropometric parameters (Fairclough, 2003; Li & Rukavina, 2012). Physical fitness is directly related to the classification of young people as active or inactive in terms of physical activity. Those who are considered inactive are significantly further away from the minimum thresholds of moderate or vigorous physical activity (Fairclough & Stratton, 2005; Martins et al., 2018). For the most part, the type of experience gained at earlier ages is the main reason why an adolescent belongs to one lifestyle or another (Li & Rukavina, 2012).

To conclude this point, it is necessary to talk about the socio-economic profile of the adolescent's family. Research supports that young people with a low socioeconomic status are often associated with lower levels of physical activity (Christofaro et al., 2018; Cvetković et al., 2014; Elhakeem et al., 2015; Molina-García et al., 2016; Yamakita et al., 2020). Based on such information, several studies have been able to argue that these values were maintained during the subjects' adulthood (Elhakeem et al., 2015; Juneau et al., 2015).

The socio-economic variable of adolescents can be linked to the influence of parents on adolescents. Parents who have had few educational opportunities tend to be less aware of the importance of physical activity for their children's health. This leads to low participation in physical activity by their offspring and, in general, to a continuation of behaviours adverse to healthy habits (Cho & Lee, 2017; Cvetković et al., 2014). Several articles have corroborated the hypothesis that adolescents with physically active parents reflect a greater willingness to engage in physical activity (Christofaro et al., 2018; Marques et al., 2017; Mitchell et al., 2012). This is driven by imitation, support and socialisation factors between children and parents (Christofaro et al., 2018; Mitchell et al., 2012; Stalsberg & Pedersen, 2010).

Pedagogical factors

Sessions devoted to activities related to an artistic profile, such as gymnastics and dance, have been associated with a significantly lower percentage of time devoted to vigorous physical activity than those focused on fitness. Regarding moderate or vigorous physical activity, it has been concluded that sessions focused on ball games obtain higher levels than the rest of the sessions (Delextrat et al., 2020). Lessons focused on invasion and fitness games also obtained high levels of moderate or vigorous physical activity time in the whole session, but below the established minimums (Beale et al., 2021; Molina-García et al., 2016).

Regarding class size, many teachers come to consider it as a major barrier to achieving the objectives set (Barroso et al., 2005).

Finally, pedagogical factors also include the variable of the space in which the session is held. Most articles have concluded that the highest values of moderate or vigorous physical activity are obtained in outdoor

spaces. This may be directly related to the type of activities performed in each space. However, it is also directly linked to institutional factors, as most educational institutions have certain limitations in terms of infrastructure (Delextrat et al., 2020; Gill et al., 2016; McKenzie et al., 2006; Molina-García et al., 2016).

Institutional factors

Within the institutional factors, curricular and governmental values can be highlighted. At this point it is necessary to emphasise the number of hours devoted to PE. Although this fact is not significant in terms of the total time of moderate or vigorous physical activity spent in a session, it is necessary to take it into account (Groffik et al., 2020). It is important to highlight the importance the subject has within the healthy habits of adolescents and to emphasise its main role, both inside and outside the school (Haerens et al., 2010; Romero-Chouza et al., 2021). In relation to this point, PE teachers should implement initiatives to be able to dedicate more minutes to activities (Dudley et al., 2012; McKenzie et al., 2000; Molina-García et al., 2016).

The direct relationship between the relevance of the problem and healthy habits means that there is an interest in the search for a solution or improvement. It is necessary to investigate the different values that may affect this fact and the difficulties that PE teachers may encounter in making young people aware of the impact that practising physical activity will have on their lives. This systematic review aims to analyse different studies that have assessed the different variables which may influence the health-related goals of PE professionals. Therefore, the main objective of this study has been to review the scientific literature to delve into the influence of secondary school PE sessions on the physical activity levels of adolescents. Additionally, as specific objectives:

- To identify whether the average levels of moderate or vigorous physical activity of secondary school students in PE sessions reach the minimum recommended levels.
- To investigate the factors that interfere with moderate to vigorous activity levels during secondary school PE sessions.

Method

Research methodology

This systematic review was conducted following the guidelines established by *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) (Page et al., 2021) and the protocols for conducting systematic reviews (Mother et al., 2015).

Search strategy

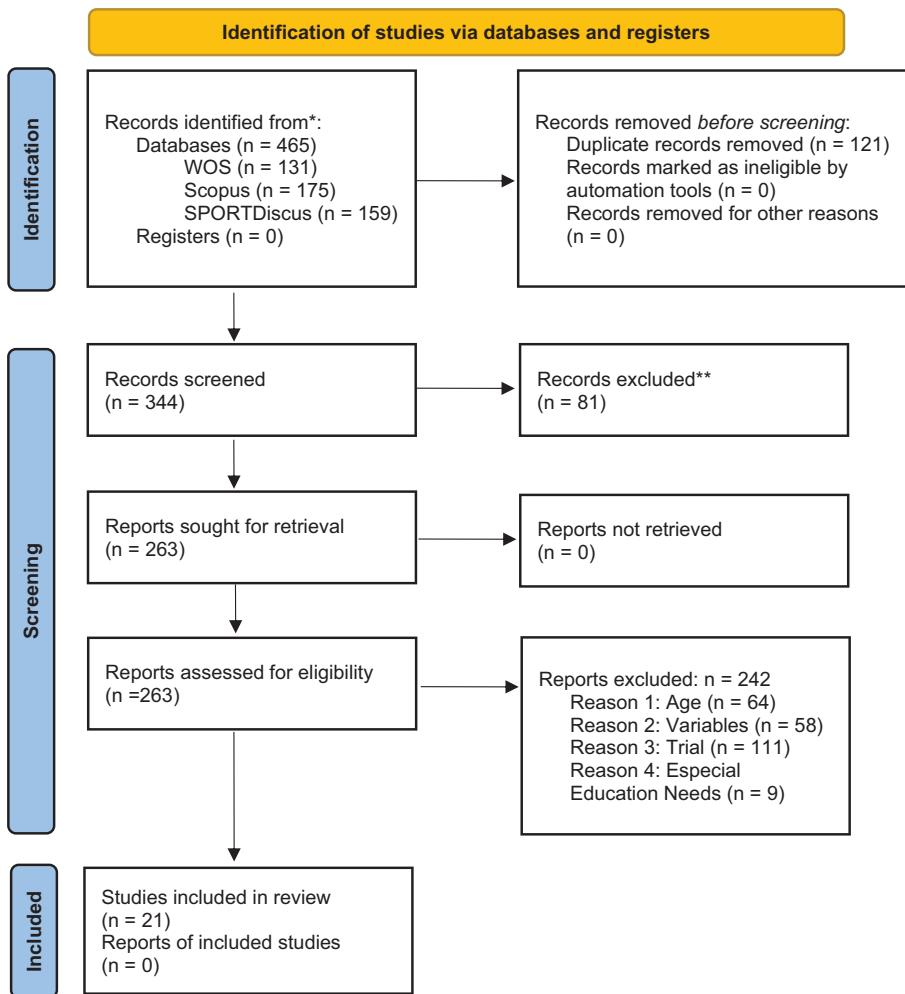
The search in this systematic review focused on three electronic databases, SPORTDiscus, Scopus and Web of Science. The search terms used were physical education, adolescent, youth, secondary, moderate to vigorous physical activity (MVPA) and physical activity level. To relate them to one another, the Boolean operators AND and OR combined with parentheses were used. For terms composed of more than one word to be considered as a single word, inverted commas were added. The final search equation is shown below: “Physical education” AND (adolescent* OR youth OR secondary) AND (MVPA OR “physical activity level”)

Eligibility criteria

In the first place, full articles in English or Spanish from 2012 to 2021 were included. Those that did not focus on subjects aged 12-18 were excluded with the intention of covering only adolescent samples.

Secondly, articles that considered variables related to the level of physical activity during PE lessons were considered valid, while those that did not have a directly related objective were considered dispensable. At the same time, those with the aim of investigating a specific methodology were excluded. After the elimination of duplicates, one author reviewed the title and abstract of the remaining papers. Both authors reviewed the full text of the selected articles. In case of disagreement, an external reviewer was used for the final decision. Finally, the systematic review focused on a total of 21 articles. Figure I shows the totals according to the database and the screening process.

FIGURE I. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only



Source: Compiled by author

Results

The 21 articles selected for the resolution of the objectives in the present systematic review were of an observational research nature. The main difference in their methodology was the type of instrument used to measure moderate to vigorous physical activity levels (12 accelerometer, 3 heart rate sensor, 2 pedometer, 2 System for Observing Fitness Instruction Time and 1 questionnaire).

TABLE I. Summary of the different articles analysed for this systematic review.

Authors and year	Type of study	Instrument	Sample	Conclusions	Variables
Aljuhani & Sand-ercock, 2019	OBS QT	ACCEL	111 adolescents aged 12-14 from four schools in Saudi Arabia.	PE lessons should be taught daily and teachers must be properly trained to deliver quality sessions.	MVPA with/without PE Physical condition
Azlan et al., 2021	OBS TR QT	ACCEL	56 students (32 boys and 24 girls) aged 13-14 from a school in Malaysia.	Playing fun activities in PE sessions can be a strategy for the promotion of physical activity during school hours. Future studies should examine what other types of games may help to increase the level of physical activity.	MVPA Contents Gender
Brus-seau & Burns, 2015	OBS QT	Pedometer ACCEL	232 students (144 boys and 88 girls) aged 12-13 in a US school.	Education practitioners should be aware of the contributions of different content regarding MVPA and step counting to provide researchers with a guide for future studies	MVPA Contents Location
Burns et al., 2015	OBS LO QT	Pedometer	100 students (62 boys and 38 girls, 62 boys) aged 12-14 from a school in the USA.	The decline during the PE course of the MVPA is leading practitioners away from one of their goals. More effective teaching methods should be applied and research into the variables affecting this downward curve should be carried out.	MVPA Physical condition Gender
Comte et al., 2015	OBS QT	Analysis of data from another study ACCEL	508 adolescents (259 boys and 249 girls) aged 16-17 in Canada, 338 did and 170 did not do PE.	The extracted data suggest that secondary school students' participation in PE can lead to improvements in physical activity levels.	MVPA Gender

(Continued)

TABLE I. Summary of the different articles analysed for this systematic review (Continued)

Authors and year	Type of study	Instrument	Sample	Conclusions	Variables
da Costa et al., 2019	OBS	Analysis of data from another study ACCEL	567 students (267 boys and 300 girls) aged 7 to 18 from two schools in Brazil.	PE professionals should raise pupils' awareness of the importance of physical activity. At the same time, motivational activities should be implemented as a strategy for increasing MVPA in schools, both in PE and in recess time.	MVPA Gender Age Physical condition Attitude NEE
Delex-trat et al., 2020	OBS QT	ACCEL	307 pupils (201 boys and 106 girls) aged 12-13 from six schools in England.	In planning PE lessons, it will be necessary to consider the location and the type of activity to be performed. Similar studies with a larger number of subjects would be helpful.	MVPA Gender Location Content
Ferreira et al., 2014	OBS QT	ACCEL	191 adolescents (98 boys and 93 girls) aged 12-17 from three schools in Portugal.	The minimum levels of MVPA during PE lessons are not reached. The results obtained serve to highlight the work that remains to be done within education in terms of strategies for increasing MVPA.	MVPA Gender Age
Hobbs et al., 2015	OBS QT	ACCEL	55 girls aged 13-14 from a school in England.	PE is not helping adolescents in the way it was intended to, teachers and education professionals must maximise PE opportunities with a focus on health.	MVPA Content Class-room contexts
Kwon et al., 2020	OBS TR QT	SOFIT	2063 sessions from 40 US schools.	Middle schools have the lowest percentage of time spent on MVPA, and more research should be done at these ages to find out how to deal with the changes that occur to improve MVPA levels.	MVPA Content SES Location
Lyrya et al., 2017	OBS QT	HR sensor	821 students (408 boys and 413 girls) aged 13-16 in 14 schools Finland.	The differences between genders are given by the content to which they are exposed. Teachers need to be aware of the benefits that each type of content promotes and plan with this knowledge while addressing the other objectives.	MVPA Gender Content
Mayor-ga-Vega et al., 2020	OBS QT	ACCEL	122 students (71 boys and 51 girls) aged 12-15 from four schools in Chile	Young people do not comply with MVPA recommendations during PE sessions. Strategies should be sought to increase the time and meet the minimums.	MVPA Gender

TABLE I. Summary of the different articles analysed for this systematic review (Continued)

Authors and year	Type of study	Instrument	Sample	Conclusions	Variables
Mayorga-Vega et al., 2017	OBS QT	ACCEL	89 students (55 boys and 34 girls) aged 13-14, from four schools in Chile.	The results should serve to alert about the need for change in educational policies with an increase in hours and professionals in PE. These changes should serve to reach the minimums established within the sessions and for the promotion of physical activity in the rest of the contexts.	MVPA Gender Physical condition Classroom context
Molina-García et al., 2016	OBS QT	ACCEL	189 students (74 boys and 115 girls) aged 16 from nine secondary schools in Spain.	Institutions involved in education should make plans to increase levels of MVPA during lessons, based on gender, personal and environmental factors.	MVPA Gender Physical condition Number of students Location SES Content
Murillo et al., 2014	OBS LO QT	ACCEL	20 adolescents (12 boys and 8 girls) aged between 12 and 14 from a school in Spain.	The differences found show the importance of the contextual level and point to the need to implement gender-tailored interventions to increase the perception of competence and thus adherence to physical activity.	MVPA Gender Content
Sanz-Martín et al., 2021	OBS QL	Questionnaire	694 pupils (364 boys and 330 girls) aged 12 to 16 from Spain	PE plays a key role in adolescents' daily physical activity levels. Different strategies should be considered to improve physical activity rates in all contexts.	MVPA With/without PE Gender
Sutherland et al., 2016	OBS TR QT	SOFIT	100 PE lessons from 10 schools in 10 areas of Australia.	Physical activity levels in disadvantaged secondary schools are below the recommended levels. Improving the quality of PE should be a priority in order to achieve one of its main objectives.	MVPA Teacher SES School
Viciiana et al., 2016	OBS QT	ACCEL	231 adolescents (120 boys and 111 girls) aged between 12 and 15 from a school in Spain.	Students do not comply with the daily MVPA recommendations; more research should be done and strategies for different contexts should be provided.	MVPA Gender Age Physical condition Classroom context

(Continued)

TABLE I. Summary of the different articles analysed for this systematic review (Continued)

Authors and year	Type of study	Instrument	Sample	Conclusions	Variables
Williams & Hannon, 2018	OBS quasi-EX QT	ACCEL	446 (240 boys and 206 girls) aged 11-16 from a school in the USA.	The effects of different groupings on the level of physical activity should be further analysed, but other factors related to motivation and attitude towards PE need to be assessed.	MVPA Gender Age
Yuste et al., 2015	OBS QT	HR sensor	182 adolescents (97 boys and 85 girls) aged between 12 and 18, from five schools in Spain.	The sessions do not meet the recommended levels of intensity, and content should be selected that involves a greater physiological involvement of the students.	MVPA Gender Content
Yuste et al., 2013	OBS QT	HR sensor	107 adolescents (60 boys and 47 girls) aged between 12 and 18 from five schools in Spain.	Although a higher level of MVPA is observed in team sports, PE classes do not meet the established recommendations. There is a need to analyse and select activities that involve higher intensity.	MVPA Gender Content

Note: MVPA – Moderate or vigorous physical activity; OBS – observation; ACCEL – accelerometer; TR – transversal; LO – longitudinal; EX – experimental; QL – qualitative; QT – quantitative; SOFIT – system for observing fitness instruction time; SES – socio-economic status; PE – Physical Education; HR – Heart Rate.

Source: Compiled by author

The levels of moderate or vigorous physical activity (MVPA) were the main variable analysed for this review. With respect to this, a difference between the percentages obtained has become apparent. Some authors argue that less than 25% of total lesson time is reached (Aljuhani & Sandercock, 2019; da Costa et al., 2019; Delextrat et al., 2020; Hobbs et al., 2015; Mayorga-Vega et al., 2017, 2020; Molina-García et al., 2016; Yuste et al., 2015), while other articles have placed the results achieved closer to the recommended minima (Lyyra et al., 2017; Murillo et al., 2014; Sutherland et al., 2016; Viciana et al., 2016). Three of the studies analysed compared the mean percentages of moderate or vigorous physical activity on days with PE and days without. In two of them, significant differences in compliance with the daily physical activity minimums were found (Aljuhani & Sandercock, 2019; Sanz-Martín et al., 2021). In another study, no significant results were obtained, but it was highlighted that PE contributed to an increase in moderate or vigorous physical activity (Comte et al., 2015).

The variable that was used the most in the studies analysed and that showed the greatest influence on physical activity levels was gender. The studies have shown that boys obtain higher percentages of moderate or vigorous physical activity than girls and, consequently, manage to come closer to the established minimums, both in PE lessons and in the whole day (Azlan et al., 2021; Comte et al., 2015; da Costa et al., 2019; Ferreira et al., 2014; Lyyra et al., 2017; Mayorga-Vega et al., 2017, 2020; Molina-García et al., 2016; Murillo et al., 2014; Viciiana et al., 2016). Two studies by the same authors obtained inverse results, with girls spending 23.47% of the total PE class in moderate or vigorous physical activity and boys 19.99% (Yuste et al., 2013, 2015). The contradiction may go hand in hand with the type of instrument used in the studies, where a heart rate sensor was used in the two latter studies and accelerometers, pedometers or other observation instruments were used in the others.

The second variable with the greatest influence on physical activity levels was the content of the sessions. The articles that have taken invasive sports into account agree that they are those that provide higher levels of moderate or vigorous physical activity during PE sessions (Brusseau & Burns, 2015; Delextrat et al., 2020; Hobbs et al., 2015; Lyyra et al., 2017; Molina-García et al., 2016; Murillo et al., 2014; Yuste et al., 2013, 2015). On the other hand, the lessons aimed at corporal expression have been found to have the lowest percentage (Delextrat et al., 2020; Lyyra et al., 2017; Molina-García et al., 2016; Murillo et al., 2014; Yuste et al., 2015). Within fitness, it has been observed that sessions in which strength was worked on have achieved more minutes of moderate or vigorous intensity (Molina-García et al., 2016). Network sports and games have not shown significant differences, but differences were observed with free-choice sessions being less productive than the previous ones (Azlan et al., 2021; Delextrat et al., 2020). The content of the PE sessions seems to be closely related to the gender variable, since invasive sports have the greatest effect on moderate or vigorous physical activity levels, and also have a greater perception of competence (Murillo et al., 2014). Consequently, girls move further away from the recommended percentages during PE sessions that address this content (Azlan et al., 2021; Comte et al., 2015; da Costa et al., 2019; Ferreira et al., 2014; Lyyra et al., 2017; Mayorga-Vega et al., 2017, 2020; Molina-García et al., 2016; Murillo et al., 2014; Viciiana et al., 2016).

In addition to gender and content, there have been less common variables across studies such as the relationship between levels of moderate or vigorous physical activity and its positive relationship with the performance of outdoor sessions (Delextrat et al., 2020; Molina-García et al., 2016). One of the studies conducted in the USA compared different school levels and found no significant differences between indoor and outdoor sessions in adolescents over 14 years of age (Kwon et al., 2020). Socioeconomic status also showed variations in intensities, with those with a lower socioeconomic status also having lower percentages of moderate or vigorous physical activity (Kwon et al., 2020; Molina-García et al., 2016; Sutherland et al., 2016). It is also important to note that three of the selected articles indicated that the older the adolescents were, the further away they moved from the established PE recommendations (da Costa et al., 2019; Ferreira et al., 2014; Williams & Hannon, 2018).

Considering the articles analysed, it can be seen that there are some differences between the different countries from which results were obtained. The USA (49%), Finland (41.14%) and Australia (39%) have obtained the highest percentages, but it is important to emphasise that only one article per country has been sampled (Kwon et al., 2020; Lyyra et al., 2017; Sutherland et al., 2016). The country with the lowest percentage of moderate or vigorous physical activity with an average of 11.83% is Chile (Mayorga-Vega et al., 2017, 2020). It can be concluded that Spain (23.7%) is at a mid-point compared to the rest of the countries; it could be placed at the same level as England and Saudi Arabia (Aljuhani & Sandercock, 2019; Delextrat et al., 2020; Hobbs et al., 2015; Molina-García et al., 2016; Yuste et al., 2013, 2015).

Discussion

The results indicate that PE professionals do not achieve the minimum of 50% of session time in moderate or vigorous physical activity set by the Association for Physical Education (2015). Therefore, adolescents' physical activity levels are insufficient to achieve health benefits (Oliveras et al., 2015). It should be noted that PE remains a facilitating channel to achieve the recommendations of the World Health Organization (2010). Likewise, it has been found that within the different contexts

of adolescents' daily lives, after-school time is the moment when most physical activity is practised. If the total time available in each setting is taken into account, PE achieves higher percentages than recess time and free time after school (Mayorga-Vega et al., 2017; Viciña et al., 2016).

In order to address this problem, it is necessary to ascertain the factors that interfere with the levels of moderate or vigorous physical activity of adolescents during PE sessions. Several studies consider that the identification of factors influencing physical activity levels is an important prerequisite for planning and developing effective educational programmes (Biddle et al., 2004; Sterd et al., 2014).

The analysis of the results has positioned gender as the variable that seems to influence pupils' physical activity levels the most. Male pupils are closer to the recommended levels of physical activity than female pupils. These data coincide with several studies that look into the practice of physical activity in relation to gender (Alvariñas-Villaverde & Pazos-González, 2018; Fernández-Villarino et al., 2019; Marques & Carreiro da Costa, 2013; Mielgo-Ayuso et al., 2016). The second variable that most seems to influence physical activity levels is directly related to gender, and that is the content of the PE sessions, as students' motivation will be different depending on their gender. Different studies on the gender perspective coincide in associating the practice of competitive sports with male students (Alvariñas-Villaverde et al., 2009) and the practice of cooperative and non-competitive physical activity with female students (Fernández-Villarino et al., 2019). In this regard, several works point to the importance of designing PE sessions that are motivating and respond to the needs of all students (Martins et al., 2017). In this way, and through student involvement, a multitude of physical activity possibilities at different levels should be offered to promote a healthy lifestyle in the long term (Durden-Myers et al., 2018). In relation to motivation, a study by Fernández-Villarino et al. (2017) suggests that engaging in physical activity outside the school environment is closely related to positive experiences in PE. In addition, Moreno-Murcia et al. (2012) consider that increased levels of physical activity are related to factors associated with personal autonomy and responsibility.

Throughout the development of this systematic review, different limitations have been observed. Firstly, the selection of articles did not take into account the instrument used during the corresponding studies. More research should have been done in this regard and an equal number of

different methods should have been obtained in order to make a more reliable comparison. Furthermore, no research was found that mentioned institutional factors. At the same time, the length of the sessions should have been taken account of, as it would have made the analysis easier if articles had been selected that dealt with the same length of time per session. What is more, a focused systematic review could have been selected to provide more extensive results on the comparison of percentages between countries. The main strength of the research is that it is the first systematic review focused on analysing the factors that influence the levels of physical activity of school-age individuals. From this point, the foundations could be laid for effective educational programmes to increase physical activity levels inside and outside the school context in order to prevent pathologies associated with a lack of physical activity and sedentary lifestyles.

Conclusions

The average levels of moderate or vigorous physical activity of the students in the PE sessions do not reach the minimum of 50% of said PE session. It has been shown that there are different factors that influence these parameters, with gender and content being the most frequently mentioned variables. Both have been directly related since invasive sports, those that provide higher levels of intensity, are the ones that motivate girls the least. There are other variables to consider such as age, location of the session and socio-economic level, among others. Finally, Spain falls within the average threshold according to the articles analysed.

In order to improve the percentages of moderate or vigorous physical activity of students in PE sessions, further research is needed to provide effective solutions. Given the importance of this subject in relation to the adolescents' lifestyles, it would be essential for the different authors to agree and carry out the research using the same method. In this way, results based on the same process would be achieved, which would facilitate comparisons between the different variables and teachers would be able to plan courses based on this information.

Bibliographical references

Aljuhani, O., & Sandercock, G. (2019). Contribution of physical education to the daily physical activity of schoolchildren in Saudi Arabia. *International Journal of Environmental Research and Public Health*, 16(13), 1-12. <https://doi.org/10.3390/ijerph16132397>

Alvariñas-Villaverde, M., Fernández-Villarino, M. A., & Villar-López, C. (2009). Actividad física y percepciones sobre deporte y género. *Revista de investigación en educación*, 6, 113-122.

Alvariñas-Villaverde, M., & Pazos-González, M. (2018). Estereotipos de género en Educación Física, una revisión centrada en el alumnado. *Revista Electrónica de Investigación Educativa*, 20(4), 154-163. <https://doi.org/10.24320/redie.2018.20.4.1840>

Association for Physical Education. (2015). *Health Position Paper*. http://www.afpe.org.uk/physical-education/wp-content/uploads/afPE_Health_Position_Paper_Web_Version.pdf

Azlan, A., Ismail, N., Fauzi, N. F. M., & Talib, R. A. (2021). Playing traditional games vs. free-play during physical education lesson to improve physical activity: a comparison study. *Pedagogy of Physical Culture and Sports*, 25(3), 178-187. <https://doi.org/10.15561/26649837.2021.0306>

Barroso, C. S., McCullum-Gómez, C., Hoelscher, D. M., Kelder, S. H., Murray, N. G., & Professor, A. (2005). Self-reported barriers to quality physical education by physical education specialists in Texas. *Journal of School Health*, 75(8), 313-319. <https://doi.org/10.1111/j.1746-1561.2005.00042.x>

Bauman, A. E. (2004). Updating the evidence that physical activity is good for health: an epidemiological review 2000-2003. *Journal of Science and Medicine in Sport*, 7(1), 6-19. [https://doi.org/10.1016/S1440-2440\(04\)80273-1](https://doi.org/10.1016/S1440-2440(04)80273-1)

Beale, N., Eldridge, E., Delextrat, A., Esser, P., Bushnell, O., Curtis, E., Wassenaar, T., Wheatley, C., Johansen-Berg, H., & Dawes, H. (2021). Exploring activity levels in physical education lessons in the UK: a cross-sectional examination of activity types and fitness levels. *BMJ Open Sport and Exercise Medicine*, 7(1), 1-9. <https://doi.org/10.1136/bmjsem-2020-000924>

Biddle, S. J. H., & Asare, M. (2011). Physical activity and mental health in children and adolescents: a review of reviews. *British*

Journal of Sports Medicine, 45(11), 886-895. <https://doi.org/10.1136/bjsports-2011-090185>

Biddle, S. J., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of sports sciences*, 22(8), 679-701. <https://doi.org/10.1080/02640410410001712412>

Brusseau, T. A., & Burns, R. D. (2015). Step count and MVPA compendium for middle school physical education activities. *Journal of Physical Education and Sport*, 15(4), 646-650. <https://doi.org/10.7752/jpes.2015.04098>

Burns, R. D., Brusseau, T. A., & Hannon, J. C. (2015). Physical activity trajectories during daily middle school physical education. *Journal of Physical Activity and Health*, 12(7), 982-989. <https://doi.org/10.1123/jpah.2014-0151>

Cantero, P., Mayor, A., Toja, B., & González, M. (2019). Fomento de estilos de vida activos en la escuela: práctica de actividad física, edad y género. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity*, 5(1), 53-69. <https://doi.org/10.17979/sportis.2019.5.1.3680>

Cho, K., & Lee, H. (2017). Effects of gender, age, race, and socioeconomic status on parents' perception of school-based health screening in the United States: Coronary Artery Risk Detection in Appalachian Communities (CARDIAC) study. *Journal of Physical Education and Sport*, 17(3), 937-941. <https://doi.org/10.7752/jpes.2017.s3143>

Christofaro, D. G., Andersen, L. B., Andrade, S. M., Barros, M. V., Saraiva, B. T., Fernandes, R. A., & Ritti-Dias, R. M. (2018). Adolescents' physical activity is associated with previous and current physical activity practice by their parents. *Jornal de Pediatria*, 94(1), 48-55. <https://doi.org/10.1016/j.jped.2017.01.007>

Comte, M., Hobin, E., Manske, S., Casey, C., Griffith, J., Leggett, C., Veugelers, P., Murnaghan, D., & McGavock, J. (2015). Is the provision of physical education to senior-years students associated with greater physical activity levels? Insight into a province-wide policy. *Journal of Physical Activity and Health*, 12(5), 649-654. <https://doi.org/10.1123/jpah.2013-0197>

Cvetković, N., Nikolić, D., Pavlović, L., Djordjević, N., Golubović, M., Stamenković, S., & Veličković, M. (2014). The socio-economic status of parents and their children's sports engagement. *Physical Education and Sport*, 12(2), 179-190.

da Costa, B. G. G., da Silva, K. S., da Silva, J. A., Minatto, G., de Lima, L. R. A., & Petroski, E. L. (2019). Sociodemographic, biological, and psychosocial correlates of light- and moderate-to-vigorous-intensity physical activity during school time, recesses, and physical education classes. *Journal of Sport and Health Science*, 8(2), 177–182. <https://doi.org/10.1016/j.jshs.2017.05.002>

Delextrat, A., Esser, P., Beale, N., Bozon, F., Eldridge, E., Izadi, H., Johansen-Berg, H., Wheatley, C., & Dawes, H. (2020). Effects of gender, activity type, class location and class composition on physical activity levels experienced during physical education classes in British secondary schools: a pilot cross-sectional study. *BMC Public Health*, 20(1), 1-12. <https://doi.org/10.1186/s12889-020-09698-y>

Dudley, D. A., Okely, A. D., Cotton, W. G., Pearson, P., & Caputi, P. (2012). Physical activity levels and movement skill instruction in secondary school physical education. *Journal of Science and Medicine in Sport*, 15(3), 231-237. <https://doi.org/10.1016/j.jsams.2011.10.005>

Durden-Myers, E. J., Green, N. R., & Whitehead, M. E. (2018). Implications for promoting physical literacy. *Journal of Teaching in Physical Education*, 37(3), 262-271. <https://doi.org/10.1123/jtpe.2018-0131>

Elhakeem, A., Cooper, R., Bann, D., & Hardy, R. (2015). Childhood socio-economic position and adult leisure-time physical activity: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1), 1-27. <https://doi.org/10.1186/s12966-015-0250-0>

Fairclough, S. J. (2003). Girls' physical activity during high school physical education: influences of body composition and cardiorespiratory fitness. *Journal of Teaching in Physical Education*, 22(4), 382-395. <https://doi.org/10.1123/jtpe.22.4.382>

Fairclough, S. J., & Stratton, G. (2005). "Physical education makes you fit and healthy". Physical education's contribution to young people's physical activity levels. *Health Education Research*, 20(1), 14–23. <https://doi.org/10.1093/her/cyg101>

Ferreira, F. S., Mota, J., & Duarte, J. A. (2014). Patterns of physical activity in Portuguese adolescents. Evaluation during physical education classes through accelerometry. *Archives of Exercise in Health and Disease*, 4(2), 280-285. <https://doi.org/10.5628/aehd.v4i2.135>

Fernández-Villarino, M. Á., González-Valeiro, M., Toja, M. B., & Carreiro da Costa, F. (2017). Valoración de la escuela y la Educación Física y su relación con la práctica de actividad física de los escolares. *Retos: nuevas tendencias en educación física, deporte y recreación*, 31, 312-315.

Fernández-Villarino, M. Á., Hernaiz-Sánchez, A., Villaverde, M. A., & Pérez-Ferreirós, A. (2019). Gender stereotypes in sport. The socialising role of the family. *Egitania Scienza*, 24, 125-135.

Gill, M., Chan-Golston, A. M., Rice, L. N., Cole, B. L., Koniak-Griffin, D., & Prelip, M. L. (2016). Consistency of moderate to vigorous physical activity in middle school physical education. *Family and Community Health*, 39(4), 283-292. <https://doi.org/10.1097/FCH.0000000000000115>

Groffik, D., Mitáš, J., Jakubec, L., Svozil, Z., & Frömel, K. (2020). Adolescents' physical activity in education systems varying in the number of weekly physical education lessons. *Research Quarterly for Exercise and Sport*, 91(4), 551-561. <https://doi.org/10.1080/02701367.2019.1688754>

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. *The Lancet Global Health*, 6(10), 1077-1086. [https://doi.org/10.1016/S2214-109X\(18\)30357-7](https://doi.org/10.1016/S2214-109X(18)30357-7)

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child and Adolescent Health*, 4(1), 23-35. [https://doi.org/10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2)

Haerens, L., Kirk, D., Cardon, G., de Bourdeaudhuij, I., & Vansteenkiste, M. (2010). Motivational profiles for secondary school physical education and its relationship to the adoption of a physically active lifestyle among university students. *European Physical Education Review*, 16(2), 117-139. <https://doi.org/10.1177/1356336X103813>

Hobbs, M., Daly-Smith, A., Morley, D., & McKenna, J. (2015). A case study objectively assessing female physical activity levels within the National Curriculum for Physical Education. *European Physical Education Review*, 21(2), 149-161. <https://doi.org/10.1177/1356336X14555296>

Juneau, C. E., Benmarhnia, T., Poulin, A. A., Côté, S., & Potvin, L. (2015). Socioeconomic position during childhood and physical activity during adulthood: a systematic review. *International Journal of Public Health*, 60(7), 799-813. <https://doi.org/10.1007/s00038-015-0710-y>

Kwon, S., Welch, S., & Mason, M. (2020). Physical education environment and student physical activity levels in low-income communities. *BMC Public Health*, 20(1), 1-9. <https://doi.org/10.1186/s12889-020-8278-8>

Li, W., & Rukavina, P. (2012). Including overweight or obese students in physical education: a social ecological constraint model. *Research Quarterly for Exercise and Sport*, 83(4), 570-578. <https://doi.org/10.1080/02701367.2012.10599254>

Lyyra, N., Heikinaho-Johansson, P., & Lyyra, M. (2017). Exploring in-class physical activity levels during physical education lessons in Finland. *Journal of Physical Education and Sport*, 17(2), 815-820. <https://doi.org/10.7752/jpes.2017.02124>

Marques, A., & Carreiro da Costa, F. (2013). Levels of physical activity of urban adolescents according to age and gender. *International Journal of Sports Science*, 3(1), 23-27.

Marques, A., González Valeiro, M., Martins, J., Fernández-Villarino, M. A., & Carreiro Da Costa, F. (2017). Relación entre la actividad física de los adolescentes y la de madres/ padres. *Revista de Psicología Del Deporte*, 26(1), 145-156.

Martins, J., Marques, A., Peralta, M., Palmeira, A., & Carreiro Da Costa, F. (2017). Correlates of physical activity in young people: a narrative review of reviews. Implications for physical education based on a socio-ecological approach. *Retos*, 31, 292-299. <https://doi.org/10.47197/retos.v0i31.53505>

Martins, J., Marques, A., Rodrigues, A., Sarmento, H., Onofre, M., & Carreiro da Costa, F. (2018). Exploring the perspectives of physically active and inactive adolescents: how does physical education influence their lifestyles? *Sport, Education and Society*, 23(5), 505-519. <https://doi.org/10.1080/13573322.2016.1229290>

Mayorga-Vega, D., Martínez-Baena, A., & Viciana, J. (2018). Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *Journal of Sports Sciences*, 36(17), 1913-1922. <https://doi.org/10.1080/02640414.2018.1425967>

Mayorga-Vega, D., Parra Saldías, M., & Viciana, J. (2017). Comparaison of moderate-to-vigorous physical activity levels between physical education, school recess and after-school time in seconday school students: an accelerometer-based study. *Kinesiology*, 49(2), 242-251. <https://doi.org/10.26582/k.49.2.1>

Mayorga-Vega, D., Parra Saldías, M., & Viciana, J. (2020). Niveles objetivos de actividad física durante las clases de EF en estudiantes chilenos usando acelerometría. *Retos*, 37, 123-128. <https://doi.org/10.47197/retos.v37i37.69238>

McKenzie, T. L., Catellier, D. J., Conway, T., Lytle, L. A., Grieser, M., Webber, L. A., Pratt, C. A., & Elder, J. P. (2006). Girls' activity levels and lesson contexts in middle school PE: TAAG baseline. *Medicine and Science in Sports and Exercise*, 38(7), 1229-1235. <https://doi.org/10.1249/01.mss.0000227307.34149.f3>

McKenzie, T. L., Marshall, S. J., Sallis, J. F., & Conway, T. L. (2000). Student activity levels, lesson context, and teacher behavior during middle school physical education. *Research Quarterly for Exercise and Sport*, 71(3), 249-259. <https://doi.org/10.1080/02701367.2000.10608905>

Mielgo-Ayuso, J., Aparicio-Ugarriza, R., Castillo, A., Ruiz, E., Ávila, J. M., Aranceta-Batrina, J., Gil, A., Ortega, R., Serra-Majem, L., Varela-Moreiras, G., & González-Gross, M. (2016). Physical activity patterns of the Spanish population are mostly determined by sex and age: findings in the ANIBES study. *PloS one*, 11(2), e0149969

Mitchell, J., Skouteris, H., McCabe, M., Ricciardelli, L. A., Milgrom, J., Baur, L. A., Fuller-Tyszkiewicz, M., & Dwyer, G. (2012). Physical activity in young children: A systematic review of parental influences. *Early Child Development and Care*, 182(11), 1411-1437. <https://doi.org/10.1080/03004430.2011.619658>

Molina-García, J., Queralt, A., Estevan, I., & Sallis, J. F. (2016). Ecological correlates of Spanish adolescents' physical activity during physical education classes. *European Physical Education Review*, 22(4), 479-489. <https://doi.org/10.1177/1356336X15623494>

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, 4(1), 1-9. <https://doi.org/10.1186/2046-4053-4-1>

Moreno-Murcia, J. A., Huéscar, E., & Cervelló, E. (2012). Prediction of adolescents doing physical activity after completing secondary education. *The Spanish journal of psychology*, 15(1), 90-100. https://doi.org/10.5209/rev_SJOP.2012.v15.n1.37288

Murillo, B., Julián, J. A., García-González, L., Abarca-Sos, A., & Zaragoza, J. (2014). Influencia del género y de los contenidos sobre la actividad física y la percepción de competencia en EF. *RICYDE: Revista Internacional de Ciencias Del Deporte*, 10(36), 131-143. <https://doi.org/10.5232/ricyde2014.03604>

Olivares, P. R., Cossio-Bolaños, M. A., Gómez-Campos, R., Almonacid-Fierro, A., & García-Rubio, J. (2015). Influence of parents and physical

education teachers in adolescent physical activity. *International Journal of Clinical and Health Psychology*, 15(2), 113-120. <https://doi.org/10.1016/j.ijchp.2015.01.002>

Organización Mundial de la Salud. (2010). *Recomendaciones mundiales sobre actividad física para la salud*. Organización Mundial de la Salud. https://apps.who.int/iris/bitstream/handle/10665/44441/9789243599977_spa.pdf

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic reviews*, 10(1), 1-11. <https://doi.org/10.1186/S13643-021-01626-4>

Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial Del Estado*, núm. 3, pp. 169-546.

Romero-Chouza, Ó., Lago-Ballesteros, J., Toja- Reboreda, B., & González-Valeiro, M. (2021). Propósitos de la EF en educación secundaria: revisión bibliográfica. *Retos*, 40, 305-316. <https://doi.org/10.47197/retos.v1i40.80843>

Sanz-Martín, D., Ruiz-Tendero, G., & Fernández-García, E. (2021). Contribution of physical education classes to daily physical activity levels of adolescents. *Physical Activity Review*, 9(2), 18-26. <https://doi.org/10.16926/par.2021.09.18>

Stalsberg, R., & Pedersen, A. V. (2010). Effects of socioeconomic status on the physical activity in adolescents: a systematic review of the evidence. *Scandinavian Journal of Medicine and Science in Sports*, 20(3), 368-383. <https://doi.org/10.1111/j.1600-0838.2009.01047.x>

Sterdt, E., Liersch, S., & Walter, U. (2014). Correlates of physical activity of children and adolescents: A systematic review of reviews. *Health Education Journal*, 73(1), 72-89. <https://doi.org/10.1177/0017896912469578>

Sutherland, R., Campbell, E., Lubans, D. R., Morgan, P. J., Okely, A. D., Nathan, N., Gillham, K., Lecathelinais, C., & Wiggers, J. (2016). Physical education in secondary schools located in low-income communities: physical activity levels, lesson context and teacher interaction. *Journal of Science and Medicine in Sport*, 19(2), 135-141. <https://doi.org/10.1016/j.jsams.2014.12.003>

Viciiana, J., Mayorga-Vega, D., & Martínez-Baena, A. (2016). Moderate-to-vigorous physical activity levels in physical education, school recess,

and after-school time: Influence of gender, age, and weight status. *Journal of Physical Activity and Health*, 13(10), 1117-1123. <https://doi.org/10.1123/jpah.2015-0537>

Williams, S. M., & Hannon, J. C. (2018). Physical Activity Levels in Coed and Same-Sex Physical Education Using the Tactical Games Model. *The Physical Educator*, 75(3), 525-545. <https://doi.org/10.18666/tpe-2018-v75-i3-8103>

Whitehead, M. (Ed.). (2010). *Physical literacy: Throughout the lifecourse*. Routledge

Yamakita, M., Kanamori, S., Kondo, N., Ashida, T., Fujiwara, T., Tsuji, T., & Kondo, K. (2020). Association between childhood socioeconomic position and sports group participation among Japanese older adults: a cross-sectional study from the JAGES 2010 survey. *Preventive Medicine Reports*, 18, 1-7. <https://doi.org/10.1016/j.pmedr.2020.101065>

Yuste, J. L., García-Jiménez, J. V., & García-Pellicer, J. J. (2015). Intensity of physical education classes in adolescents. *Revista Internacional de Medicina y Ciencias de La Actividad Física y Del Deporte*, 15(58), 309-323. <https://doi.org/10.15366/rimcaf2015.58.007>

Yuste, J. L., García-Jiménez, J. V., García-Pellicer, J. J., Vicente, J., & Jiménez, G. (2013). Intensity of physical education classes: team sports vs. individual sports. *Cultura, Ciencia y Deporte*, 8(24), 183-190. <https://doi.org/10.12800/ccd.v8i24.356>

Zhou, Y., & Wang, L. (2019). Correlates of physical activity of students in secondary school physical education: a systematic review of literature. *BioMed Research International*, 2019, 1-12. <https://doi.org/10.1155/2019/4563484>

Contact address: Ariadna Hernaiz-Sánchez. Universidad Europea de Madrid, Facultad de Ciencias Sociales y de la Educación, Departamento de Educación y Humanidades. c/Tajo, s/n. 28670 Villaviciosa de Odón, Madrid. E-mail: ariadna.hernaiz@universidadeuropea.es

Differential factors in Vocational Education and Training itineraries: A longitudinal study

Factores diferenciales en los itinerarios en Formación Profesional: un estudio longitudinal

<https://doi.org/10.4438/1988-592X-RE-2023-401-587>

Míriam Abiétar López

<https://orcid.org/0000-0003-2082-2434>

Universitat de València

Joan Carles Bernad i Garcia

<https://orcid.org/0000-0002-9376-1225>

Universitat de València

Ana Isabel Córdoba Iñesta

<https://orcid.org/0000-0001-9855-1174>

Universitat de València

Elena Giménez Urraco

<https://orcid.org/0000-0001-7092-3669>

Universitat de València

Esperanza Meri Crespo

<https://orcid.org/0000-0002-5150-1919>

Universitat de València

Almudena Adelaida Navas Saurin

<https://orcid.org/0000-0002-2301-6932>

Universitat de València

Abstract

The study presented here is part of a national research project, which was then continued at a regional level, focusing on the pathways of students at the initial level of Vocational Training (Basic Vocational Education and Training, hereinafter BVET) and the intermediate level (Intermediate Vocational Education and Training, IVET). In this article we present a longitudinal study of the pathways of VET students in the region of Valencia, which aims to identify the variables that make the difference between those who have completed their studies and those who are repeating them or have dropped out, considering socio-demographic variables, variables on their previous pathway, as well as the different dimensions of engagement. The research was conducted by passing questionnaires over three academic years, starting in 2016-2017. The results have been analysed using a Generalised Linear Model for the whole sample and, subsequently, the analysis is also presented for students at both levels, with the aim of explaining the differences between the three proposed pathways. The results indicate that those students who feel that they work hard to do their schoolwork, who enjoy attending school, who value the good relationship they have with their peers, and who see their studies as a value for their professional future, are those who complete their VET studies, although there are some important differences in each of the subsamples. Finally, we highlight as a key factor for the completion of studies among VET students the importance of giving future meaning to the current training process.

Keywords: Vocational Education and Training, longitudinal study, pathways, engagement, early school leaving.

Resumen

El estudio que presentamos se enmarca en una investigación estatal, que tuvo continuidad a nivel autonómico, centrada en conocer los itinerarios del estudiantado del nivel inicial de Formación Profesional -FP- (Formación Profesional Básica -FPB-) y del intermedio (Formación Profesional de Grado Medio -FPGM-). En este artículo presentamos un estudio longitudinal de los itinerarios del estudiantado de FP de la provincia de Valencia, que tiene por propósito identificar las variables que nos permitan diferenciar entre aquellas personas que finalizaron sus estudios, frente a las que están repitiéndolos o los abandonaron; considerando variables sociodemográficas, variables sobre su itinerario previo, así como las diferentes dimensiones del *engagement*. La investigación se concretó en el pase de cuestionarios durante tres cursos académicos, comenzando en 2016-2017. Los resultados se han analizado mediante un Modelo Lineal Generalizado para toda la muestra y, posteriormente, se presenta también el análisis para alumnado de FPB y de FPGM, con el objetivo de poder explicar las diferencias entre los

tres itinerarios propuestos. Los resultados nos indican que aquel alumnado que siente que se esfuerza en cumplir con sus tareas escolares, que le gusta asistir al centro educativo, que valora la buena relación que mantiene con sus iguales, y que entiende sus estudios como un valor para su futuro profesional, es aquel que finaliza sus estudios de FP, aunque aparecen algunas diferencias importantes en cada una de las submuestras. Finalmente, destacamos como una clave central para la finalización de los estudios entre el alumnado de FP el dotar de sentido futuro el proceso formativo presente.

Palabras clave: Formación Profesional, estudio longitudinal, itinerarios, *engagement*, abandono educativo temprano.

Introduction

Increasing the educational level of the population is one of the key points of the dominant educational discourse: promoting inclusive, equitable, quality education that favours lifelong learning is proposed as a framework for action towards which educational policies should be orientated (UNESCO, 2015), emphasizing the relationship between the educational pathway followed and the possibilities of socio-occupational integration. From this perspective, vocational educational training (VET) stands out as a key to both personal development and the productive development of the state in terms of human capital.

Currently, the training offer of the Spanish VET system (Organic Law 3/2022 on the organisation and integration of Vocational Training) integrates the offer included in basic education, that included in VET studies of the educational system, that linked to the professional competence standards of the National Catalogue of Professional Competence Standards, and that aimed at specific groups. VET studies of the educational system include three levels: Basic Vocational Education and Training (hereinafter, BVET), Intermediate Vocational Education and Training (IVET) and Higher Vocational Education and Training (HVET). The system is organised in such a way that the initial level is accessed from the third year of Compulsory Secondary Education (CSE), a derivation that reinforces the compensatory nature of this level. Access to the intermediate level (IVET) is proposed for those who have obtained the Certificate of Compulsory Secondary Education (CCSE) or BVET certificate. In general terms, this level is proposed in parallel to the Baccalaureate studies in the transition from compulsory to post-compulsory education, thus differentiating an

academic pathway from a vocational one. Finally, access to the higher level (HVET) is established for those who have a Technical Degree, giving continuity to IVET, although it can also be accessed from other degrees, such as the Baccalaureate or a university degree.

Although these are the pathways included in the current educational system, it is important to consider the different routes that students can take, “de-standardising” the regulatory proposals that pose a linearity between compulsory and post-compulsory studies in which the academic pathway (CSE and Baccalaureate) continues to have a higher status than the vocational training pathway (IVET and HVET). As an example of this diversification, we can think of those who decide to take an IVET after accessing the Baccalaureate, those who take a HVET after completing their university studies or those who take several programmes in a process of accumulation of degrees. In view of this diversity, research on pathways allows us to understand the different ways in which students navigate the educational system beyond the regulatory proposals.

From these initial considerations, in this article we address training itineraries, understanding them as processes developed by students along the pathways established by the educational system, leading to various possibilities for professional transition and transition to adult life (Casal et al., 2006). In addition, the research we present here focuses specifically on the first two levels of VET (BVET and IVET). We approach VET as an educational context in its own right that enables the development of professional careers, despite the fact that its position in the educational system reinforces its condition as a secondary pathway, the main exponent being the basic level, aimed at those persons in a situation (or at risk) of educational exclusion; a fact that since the General Education Law (approved in 1970) has supported the subsidiary nature of vocational training (Merino, 2013).

In this sense, successive reforms of the Spanish education system, including of the VET system, have addressed comprehensiveness. The comprehensive approach in education marks a turning point and a change of course from the academicist, selective and elitist practices that characterised the educational system before and during the dark years of the dictatorship and well into the 1980s (De Puelles, 1999). This approach clearly differs from the previous ones, among other aspects, in that it defends the need to attend to the diversity contained in educational centres, reducing the rigidity of educational tools, with the evident

intention of achieving the civic ideal of education for all (Bernad & Molpeceres, 2006). But this updating of the civic commitment of education presents a very significant peculiarity; namely, it raises the need to be more flexible and to adapt the contents to the students' skills, that is, to introduce dynamics within the educational system that favour its flexibility and adaptability (Martínez et al., 2015).

In this reformist context, in which the various stages that make up the education system have been brought together, one of the debates surrounding the compulsory secondary stage is whether it should have a finalist or a propaedeutic sense (Abiétar & Navas, 2017). In this debate, the meaning of VET also comes into play, considering its position in the system as a whole. In this way, one should emphasise its comprehensive function as it is an alternative to academic education both in the compulsory and post-compulsory stages. Although it is the BVET that is explicitly linked to the compensatory programmes within the framework of the second opportunity, it is worth noting that both the IVET and the HVET provide other possible paths for students, broadening the training options and, therefore, the continuity of the itineraries.

The inclusion of VET in the secondary stage and its connections with the other pathways that make up the compulsory and post-compulsory stages have been regulated differentially, for example, by modifying access or qualification requirements. These regulatory changes in turn imply changes in the possible itineraries, as they modulate the costs and benefits of their development. In this sense, in addition to institutional resources, the biographical dimension should be taken into account (Casal et al., 2006), assuming that the construction of itineraries is influenced by the individual's own choices, framed in turn by the constraints of the environment.

Taking itineraries as the object of study, and assuming the biographical perspective as a perspective to analyse them, the longitudinal methodological approach allows us to better understand the processes followed by students in the educational institutions (Casal et al., 2011). That is to say, it makes it possible to approach the processes of construction of the formative itineraries beyond the paths established by the system.

One of the difficulties in developing this methodological approach is the availability of data to reconstruct the itineraries. Educational statistics periodically provide us with a description of the situation of the system (for example, with enrolment and graduation data), but they are not sufficient

for such a reconstruction since we do not have all the elements to allow us to match one situation with another. In this sense, as proposed by Garcia et al. (2022, p. 118), it would be necessary to complete the statistics with “an annual registry of students from the moment they begin their journey in the educational system (...) that allows us to follow and analyse their educational itineraries throughout their schooling, a registry which, in fact, already exists in some autonomous communities”.

Recent precursors of longitudinal studies in the context of Spain include the exploitation of the retrospective survey ETEFIL 2005 (Encuesta Transición, Educación, Formación e Inserción Laboral), which allowed a “longitudinal and biographical approach that reconstructs training and work itineraries” (Garcia et al., 2013, p. 65). Broadly speaking, the analysis carried out by the Grup de Recerca en Educació i Treball (GRET) showed a distribution by thirds of the itineraries: one third of the students finished CSE and left the system; another third continued with post-compulsory education and finished either with a Baccalaureate degree or with a vocational training degree (IVET or HVET) and did not continue their education; and the remaining third went on to university, thus developing a prolonged school itinerary.

Based on the same data, GRET carried out a specific analysis of the training and employment pathways of VET students (Merino et al., 2011). In this case, the terminal nature of the IVET stands out, since 75% of those who graduated at this level did not continue their training itinerary. It should be noted that these data refer to 2001, and, therefore, to a context of economic boom in which early school dropout increased significantly in favour of insertion into the labour market. However, the phenomenon of specialisation through the accumulation of VET programmes is also reported: 6% in IVET and 8% in HVET. In fact, at this last level, the analysis indicates that the extension of the itineraries was more generalised; most significantly, with 30% of the students going on to university. Thus, only 57% of the graduates entered the labour market with only one degree.

If we look at more recent data, a first analysis of the 2019 Survey of educational-training transition and labour market insertion, presented by the National Statistics Institute (INE), highlights that 45.2% of those who graduated in IVET in 2013-14 dropped out of the educational system, 10.6% continued in HVET the following year and 11.1% dropped out and later rejoined HVET (INE, 2019).

The statistics of the Ministry of Education and Vocational Training (MEFP), in which VET itineraries can be analysed to a certain extent, include both the educational continuity of those who graduate and their labour market insertion. Focusing on the first element (“Subsequent educational follow-up of Vocational Training graduates”¹) and placing ourselves in the courses closest to our research, the results relating to people who graduate from VET and, in the following courses, enrol in other studies indicate that, in the three years following graduation in 2016-2017, 64.6% of those graduating from BVET accessed IVET; 42.8% of people graduating from IVET started HVET; and 27.1% of those graduating from HVET pursued University Studies.

On the other hand, the results show that age is a relevant variable as regards to continuity in the system, so that, as it increases, the percentage of students who continue studying decreases. Likewise, the professional family is another aspect to be considered, with relevant differences in the offer as a whole; most significantly, the continuity in Graphic Arts (71.1%), Computer Science and Communications (68.9%) from BVET to IVET and in Computer Science and Communications (72.9%) and Physical and Sports Activities (68.4%) from IVET to HVET.

The MEFP also offers the academic results for each level of VET, which would allow an approximation of the cohort itineraries, although not the reconstruction of itineraries in the sense previously proposed. However, the statistics available for the province of Valencia do not allow us to track enrolment and graduation for the same academic year. Thus, the data presented here complete these statistics, in part because they do not only include the itineraries of those who graduate, but of the entire population enrolled in vocational training, including the initial level, which is left out of the official statistics.

Moreover, it should be noted that we do not approach VET from the perspective of a choice in the transition from compulsory to post-compulsory studies. Our starting point is the population already enrolled in vocational training, at the initial or intermediate level, from which we intend to study the pathways followed. Thus, the previous itinerary in the compulsory stage is one more variable to be considered in our study. Specifically, in the proposal we present with the aim of studying

¹ <https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/laborales/seguimiento.html>

VET pathways, we take into account, in addition to the previous pathway, sociodemographic variables (gender, parents' education and family socioeconomic situation) and the different dimensions of engagement (behavioural and academic, emotional and cognitive).

In the case of the previous itinerary, as detailed by García et al. (2022), there are several indicators that make it possible to measure educational traceability - understood as "the ways in which young people move toward compulsory education" (p. 102). Specifically, they highlight the relationship that previous experiences of grade repetition, as well as "situations of rupture and educational interruptions", may have with the development of "non-linear" or non-traditional trajectories, such as those characterised by situations of repetition, reorientation, discontinuity, or abandonment:

Some research studies have analysed the repetition experiences of students and their impact on academic performance, showing that students with repetition experiences have higher probability of presenting dismal results in PISA tests (Calero et al., 2010) and increasing detachment and low motivation (Krüger et al., 2015) and present attitudinal problems and school disaffection (Méndez & Cerezo, 2017). (García et al., 2022, p. 102).

With regard to sociodemographic variables, in the words of Jacovkis et al. (2020, p. 284), "the literature confirms the relevance of socioeconomic and cultural factors in the analysis of educational trajectories and questions the foundation of discourses centred on individual skills as defining elements of educational success". Many research studies have shown that it is young people with low cultural and economic capital who are most represented in school dropout pathways (García et al., 2013; Julià, 2018; Valdés, 2019; as cited in García et al., 2022). In this sense, García and Valls (2018) identify four academic itineraries, finding that young men whose parents have no education or only compulsory education are overrepresented in the dropout and repetition itineraries. In the same sense, there is abundant evidence that the probabilities of having greater school success increase with higher educational levels of the parents (Carabaña, 1993; Calero, 2007; Forquin, 1985; Jackson et al., 2007; as cited in Torrents et al., 2018). In fact, the latest report on the National System of Education Indicators (MEFP, 2021), confirms, in line with previous reports, the importance of the parents' level of education in dropout rates, especially that of the mother. Specifically, the percentage varies from 3.6% in the case of young people whose mothers have higher

education to 10.1% if they have a second stage of secondary education. The percentage rises to 20.5% for lower secondary education and 39.2% for primary education or lower. According to Boudon (1974, as cited in García et al., 2022), the effects of the social origins not only act directly on academic results, they also impact unequal individual and family aspirations and expectations, with the perception of risk and success of a given educational option and its economic and opportunity costs.

As regards engagement, we start from the proposal of Reschly and Christenson (2012), this being the theoretical reference of the research work on which the results presented are based. This research work is explained in the following section. Specifically, these authors explain that engagement “is considered multidimensional, involving aspects of students' emotion, behaviour (participation, academic learning time) and cognition” (Reschly & Christenson, 2012, p. 3). In other words, engagement does not only refer to cognitive issues, but also to other dimensions, such as the emotional and behavioural dimensions. This extends the educational process to other contexts and agents beyond the classroom and the relationship between the teacher and students, such as the family and peer relationships. In this vein, Reschly and Christenson (2012) themselves present different studies that confirm the relationship between interventions focused on fostering student engagement and the prevention of educational dropout.

From another position, the relationship between engagement and the “centre effect” is proposed, focusing on the influence of the social composition and the mechanisms for attending to diversity (Tarabini et al., 2019), i.e., referring to the dynamics of the centres and their influence on the students' engagement in their studies. Also focusing on the operating dynamics of the centres, González and Bernárdez (2019) stress the relevance of the educational experience provided by the centres themselves, addressing issues such as the climate and the relationships that are formed or the support that students may receive at various points in their educational itinerary.

Thus, engagement is a variable to be considered in the analysis of pathways insofar as it is related to the students' experience in their educational process and, specifically, to whether they continue or drop out of their studies. Considering engagement on a multidimensional level broadens the understanding of the educational experience and, therefore, the variables that can enable us to understand the construction of itineraries in a given structure.

In summary, our objective is to analyse which variables have an impact on the different pathways of success, repetition or dropout, taking into account relevant sociodemographic variables, variables related to the educational trajectory, and the different dimensions of engagement.

Method

In this article we present the analysis of the data obtained in the national research project “Itineraries of success and dropout in Vocational Training in level 1 and 2 of the education system” (EDU2013-42854-R). This research had continuity through the regional research project “Itineraries of success and dropout in Vocational Training in level 1 and 2 of the education system of the province of Valencia” (GV/2018/038), which replicated part of the national research with a year's difference. Both research works focused on studying dropout in VET and developing proposals for its prevention, intervention and remediation. They were based on a longitudinal methodology that combined several methods for data collection: statistical data analysis and longitudinal questionnaires over three years.

Sample

The sample for the design of the regional project was based on enrolment data of BVET and IVET in the province of Valencia in the academic year 2016-17, provided by the General Directorate of Vocational Training and Special Regime Studies of the Valencia Regional Government. Specifically, the sample universe was 5,288 students in BVET and 21,246 in IVET. The sample was stratified by professional family, geographic location and type of centre.

The sample for the three courses of the research is specified in the data presented in Table I. In BVET, the optimal sample design, with a confidence level of 95% and a sampling error of 3%, consisted of 894 questionnaires; while in IVET, of 1,028 questionnaires. With the questionnaires obtained in the first course, the real error of the sample was 3.35% in BVET and 2.27% in IVET.

TABLE I. Sample for the longitudinal analysis of itineraries in the province of Valencia.

	Questionnaires obtained in BVET			Questionnaires obtained in IVET		
	2016-17	2017-18	2018-19	2016-17	2017-18	2018-19
Total	739	398	411	1,234	498	676

Source: Compiled by authors

Specifying the composition of the sample obtained (described in detail in Navas et al., 2021) and considering the results presented below, it should be noted that more than 50% of the questionnaires obtained refer to the following professional families: Administration and Management, Electricity and Electronics, Healthcare, and Computer Science and Communications.

The results presented here correspond to a cross-analysis between the data from the first and third years of the research. We had responses from 1,087 subjects, 411 of whom were enrolled in the first year in BVET and 676 in GM, implying a sample loss of about 45% in both cases.

Instrument

The instrument used was designed ad-hoc for the research project “Itineraries of success and dropout in Vocational Training in level 1 and 2 of the education system”, having as a theoretical reference the concept of engagement from the proposal of Reschly and Christenson (2012), as detailed by Cerdà-Navarro et al. (2019). From this perspective, the questions included three dimensions related to engagement (behavioural and academic, emotional and cognitive), as well as four others to complement the information on the students: personal and family characteristics, previous educational pathways, occupational and training pathways and life events (Table II).

The third questionnaire also included questions relating to the current situation, both economic (economic independence, difficulty in making ends meet) and educational (situation with respect to studies started in 2016-17). In this article we focus on the latter, relating it to the previous itinerary, sociodemographic variables and engagement.

TABLE II. Dimensions and variables included in questionnaire Q1.

Behavioural and academic dimension	School effort Indiscipline School adhesion Participation in leisure and free time activities outside the centre Participation in leisure and free time activities within the centre
Emotional dimension	Relationships with teachers Relationships with peers Family support Perception of parental commitment
Cognitive dimension	Control and relevance of schoolwork Future aspirations and achievements and expectations of professional results Intrinsic motivation Professional identity
Personal and family characteristics	
Studies (previous educational itineraries)	
Occupational and training itinerary	
Life events	

Source: Compiled by authors based on Navas et al., 2021 (pp. 201-202).

Procedure

In the first two phases of the research, the questionnaire was handed over on paper in the selected centres. In the third phase, telephone interviews were conducted from the sample obtained in the first phase, which made it possible to recover subjects lost in the second year due to different reasons such as absenteeism, change of studies/centre, or dropout.

Regarding the analysis of the data obtained, three general linear models were used to determine the variables that best explained dropout, year repetition or completion of studies. The generalised linear model is a type of linear regression that allows analysis with ordinal or discontinuous

variables and determines the explanatory variables with respect to the dependent variables of the study.

The first model considered the two levels of FP, both BVET and IVET, with an Akaike index (AIC) of 1,146.06. The second focused on BVET cases and the third on IVET students. The latter two considerably improved the Akaike index with 457.53 for BVET and 717.25 for IVET. However, we will present the results of the three models in order to give a more complete picture. Another generalised linear model was performed to test whether the AIC improved when including the FP Level variable as an explanatory variable and the interactions of this variable with the rest of the explanatory variables. However, in this case the Akaike index was even higher, reaching 1,164, so this analysis was not included.

In the three models, we analysed each of the dimensions of engagement (behavioural, cognitive and affective) as explanatory variables, as well as each of the factors that compose them. Furthermore, sociodemographic variables that were considered more relevant to the research were analysed: sex, level of education of the mother and father, and whether they (the students) had economic difficulties. Another object of analysis was the number of years repeated in ESO and whether the students had been suspended or expelled from school.

As a dependent variable, we considered the possible itineraries when studying FP, where the alternatives were: dropout, repetition or completion.

Results

Generalised linear model for the two levels of VET

In the model focused on the two levels of VET, significant differences were observed in different variables. With respect to behavioural engagement, we found significant differences in school effort ($\chi^2=15.47$, $p<0.001$), where the greater the school effort, the greater the probability of completion, with a clear upward trend where dropout has an average of 1.78 (0.54) and completion of 2.14 (0.54).

With respect to the activities performed, only those conducted outside the centre were significant ($\chi^2=6.5$, $p=0.011$). Students with a dropout (2.07 (0.97)) or repetition profile in FP (2.07 (0.91)) participate in more activities outside the centre, and fewer with a completion profile perform such activities (1.92 (1.07)).

Regarding affective engagement, significant differences were only observed in relationships with peers ($\chi^2=7.19$, $p=0.007$) with a less clear trend. It is observed that better relationships favour both repetition ($\bar{x}=2.19$ (0.54)), i.e., staying at the centre with the intention of finishing, and completion ($\bar{x}=2.18$ (0.47)). The result is lower in the case of drop-out ($\bar{x}=2.15$ (0.49)).

Differences in cognitive engagement were found in Future Aspirations and Achievements ($\chi^2=4.27$, $p=0.039$). A clear downward trend is observed, so that higher aspirations mean higher probabilities of completion ($\bar{x}=2.35$ (0.48)), followed by repetition ($\bar{x}=2.24$ (0.46)) and drop-out ($\bar{x}=2.06$ (0.55)).

Differences were also found in School Adherence ($\chi^2=5.89$, $p=0.017$). In this case the results are very similar for both repetition ($\bar{x}=1.98$ (0.85)) and completion ($\bar{x}=1.94$ (0.77)), with greater adherence in both cases compared to drop-out ($\bar{x}=1.45$ (0.9)).

As for the sociodemographic variables, we only found differences in some variables. In this regard, the one referring to the mother's studies was clearly significant ($\chi^2=21.07$; $p=0.004$). There are differences to be seen between the group that repeated a grade ($\bar{x}=2.81$ (2.32)) and the rest, which shows a higher score. This would indicate that the higher the mother's level of education, the higher the probability of completion ($\bar{x}=3.35$ (2.35)) followed by drop-out ($\bar{x}=3.16$ (2.44)).

With respect to the previous pathway, the variable Number of Years Repeated in ESO is significant ($\chi^2=11.52$; $p=0.042$). A clear downward trend is observed, so that the fewer years repeated in ESO, the greater the probability of completion ($\bar{x}=0.97$ (0.85)) compared to that of repetition ($\bar{x}=1$ (0.79)) and drop-out ($\bar{x}=1.19$ (0.72)).

Finally, the fact of having been suspended or expelled from the centre was also significant ($\chi^2=4.56$; $p=0.033$). In cases of suspension there is a higher probability of abandonment (0.63 (0.48)). The other two cases have similar scores: repetition ($\bar{x}=0.84$ (0.37)) and completion ($\bar{x}=0.83$ (0.37)).

Generalised linear model for BVET students

Secondly, we performed a generalised linear model with the same variables mentioned in the previous model, exclusively for BVET students, to identify the variables that explain completion, repetition or drop-out. In

this case, we observed that there were no significant differences in any variable in any of the engagement cases. Nor were there significant differences with respect to the previous itineraries for BVET.

Generalised linear model for IVET students

Finally, we performed a generalised linear model for the IVET students. In this case, the variables that are significant correspond mostly to the significant variables of the general analysis, where both levels of FP are taken into account, observing differences at the level of behavioural engagement in school effort ($\chi^2=20.04$; $p<0.001$). There is a clear upward trend from lower to higher school effort from dropout ($\bar{x}=1.74$ (0.5)) to success ($\bar{x}=2.18$ (0.54)), as happens in the general analysis.

The variable referring to the activities that students carry out outside the centre also turned out to be significant ($\chi^2=4.75$; $p=0.029$). In this instance, a downward trend is observed, so that the more time dedicated to activities outside the centre, the lower is the degree of completion among students, with an average dropout rate of 2.14 (0.95), a repetition rate of 2 (0.91) and completion rate of 1.9 (1.05), similar to the results of the general analysis.

Regarding affective engagement, as in the case of the general analysis, differences were found according to relationships with peers ($\chi^2=15.68$; $p<0.001$), although a great variability was observed that does not appear in the general analysis; thus, very similar scores were found both in the cases of abandonment ($\bar{x}=2.22$ (0.46)) and completion ($\bar{x}=2.21$ (0.45)), with the highest score for repetition ($\bar{x}=2.26$ (0.54)), which would indicate that better relations with peers promotes cases of repetition and, to a lesser extent, of completion and dropout.

Regarding cognitive engagement, there are differences in School Adherence ($\chi^2=5.91$; $p=0.015$) in the same way as in the general analysis, although the trend is significantly modified. Here the difference is observed between dropout ($\bar{x}=1.59$ (0.82)) and the rest, so that there is clearly less adherence in the cases of abandonment compared to repetition, which has somewhat higher scores ($\bar{x}=2.07$ (0.82)) than the cases of completion ($\bar{x}=2.03$ (0.69)).

With respect to sociodemographic variables and previous itineraries, we found no significant differences.

Discussion

We now establish a dialogue between our results and the work presented in the introduction, as well as with other contributions relevant to their understanding. To do so, we will analyse results related to the sample in general and we will focus on those cases where we have found differences between the complete sample and the BVET and IVET subsamples.

With regard to the total sample, the scientific literature stresses the importance of the previous itinerary. Our results concur in highlighting past events such as repetition in high school and having been suspended or expelled from school. Both facts seem to feed the disengagement with the centre and its dynamics, favouring student dropout (García et al., 2022).

Variables that do not have a significant impact are repetition in primary school and the professional family in which the students are being trained. In the first case, the non-significance may be based on the exceptional nature of the measure in the primary stage, where teachers operate with a higher level of comprehensiveness and less rigidity regarding the achievement or non-achievement of objectives in each grade, placing value on the continuity of the class group and the relationships among its students.

In the second case, the lack of a sufficient sample in each professional family and the imbalances between them led us to group them into the three basic productive sectors, which blurs some possible differences and unites professional families as distant as, for example, Health and Administration & Management within the tertiary sector. It seems necessary to obtain a sufficient sample in each of the families to be able to study their effect, for which it would be essential to handle data on a larger territorial scale so that they could be statistically exploitable.

Regarding sociodemographic variables, our results mark a somewhat different line from what has been identified in the scientific literature (Jacovkis et al., 2020; Torrents et al., 2018). In our study, we see how variables such as sex and family economic difficulties do not discriminate between those who drop out, repeat or complete their studies; with the exception of the role of the parents' education, where there is a clear influence of the level of education attained by the mother, but not by the father (MEFP, 2021), although this only happens when the complete VET sample is taken. Another key factor, age, was not considered since most

of the groups were of the same age, a fact which became almost a constant in our research. The lack of significance of the sex variable may be due to the fact that, in general, we are working in training areas focused on applied and manual knowledge, with low academic demands, where motivation, vocation and disinterest seem to be more equally distributed among students of both sexes. Historically, the context of VET has been socially undervalued and marked by class but also gender structuring, as Martínez and Merino (2011) point out, although the class perspective has predominated when analysing the context of VET (Niemeyer & Colley, 2015). Studies in this area emphasise that divisions by social class become evident when comparing academic and career paths, while gender division is more clearly seen when confronting different career paths (Nylund et al., 2018). The lack of studies that consider the sex and/or gender variable in VET suggests the need for future research that includes this perspective as its central focus.

Also striking is the apparent independence between the economic difficulties of families and the students' academic trajectories. We believe that this hides the fact that we are dealing with a training programme mainly attended by students of similar socioeconomic levels. On the other hand, one type of cultural capital (Bourdieu, 2001), represented by the mother's level of education, maintains its discriminatory capacity, where the higher the mothers' level of education, the greater the probability of completing the studies. The lack of influence of the father's education and the significance of the mother's refer us to educational patterns centred on the traditional distribution of gender roles within the family context, where it is the woman who to a greater extent accompanies the formative process of her children. Therefore, differences in their cultural capital have a significant influence on their children's trajectories.

In relation to the role of engagement, we see that all its dimensions show the ability to discriminate in one sense or another. In the academic and behavioural dimension, both school effort and school adherence in the full sample as well as the IVET sample score significantly higher among students who completed their studies. This was not confirmed in the BVET sample. Something similar happens in the emotional dimension in terms of their assessment of the relationships they establish with their peers in class, as well as in the cognitive dimension as regards their aspirations, achievements, and expectations, although this was only significant in the full sample. Thus, in line with previous work (Reschly &

Christenson, 2012), our work confirms the importance of engagement when studying student pathways. However, our research allows us to further refine this aspect, thanks to the concrete identification of the different components of the dimensions of engagement that show discriminatory capacity, thus presenting a picture that allows us the audacity to predict that those students who feel that they are making an effort to fulfil their obligations, who like to attend school, who consider that they maintain good relationships with their peers and who understand their studies as a value for their professional future, will be those who most likely will complete their studies.

The results of two components of the academic and behavioural dimension of engagement deserve a separate mention: the significance of participation in leisure and free time activities outside the school context for both the complete sample and the IVET sample and, on the other hand, the non-significance of participation in extracurricular activities within the school. Both cases are contradictory to each other and do not show clear explanatory elements, so that in future works we shall have to fine-tune the data collection elements in order to try to explain them more clearly.

Furthermore, it is worth paying attention to the set of components of the three dimensions of engagement that do not seem to be able to make a difference as regards the different trajectories: in the academic and behavioural dimension, the component referred to would be indiscipline; in the emotional dimension, we refer both to the valuation of relationships with teachers and to family support and commitment; and, finally, in the cognitive dimension, we mean the relevance of school work, motivation and identification with the profession for which they are preparing.

Starting with indiscipline, we see that it does not behave as we would expect. We consider that the low general levels reported by students make it difficult for this variable to show any type of discriminative capacity. We would need to be able to triangulate the students' information with that of other significant agents, since we are probably faced with the fact that what the students do not consider 'indiscipline' can be understood as such by the teachers.

As for the components of the emotional dimension that do not show significance, it looks like the referents of the adult world do not play a relevant role regarding the different educational trajectories of the stu-

dents. We could say that students evolve independently of the relationships and involvement of teachers and parents, whereas their peers do play a significant role. This allows us to venture into explanatory hypotheses that take into account the central role of the peer group at these ages.

In connection with the components of the cognitive dimension that do not show significance, those that can give meaning to the training experience, such as the relevance of the activities we carry out, how motivated we are by the training we receive, or whether we feel we can be professionals in the field, do not show discriminatory capacity. It would seem as if everything related to intrinsic motivation, to the value of training in itself, does not play a relevant role as regards the likelihood of finishing, repeating or dropping out. In other words, finding meaning in training in itself does not seem to be the key. Their approach seems to be more strategic, more pragmatic: it is the students who give future meaning to their present training who complete their studies the most (as opposed to those who repeat and those who drop out), which would be expected in studies whose clear purpose is insertion into the labour world. In short, it is the students that can give a meaningful answer in their lives to the difficult question: "What's the point of all this?"

If the above shows an overall picture of the entire group surveyed, the separate analysis for each sample group also reveals some differences.

When we look exclusively at students in BVET, we find that none of the variables studied attain significance, and therefore do not allow us to distinguish between students who finish, repeat or drop out.

The fact that none of the variables discriminate sufficiently suggests, at least, a couple of issues to continue working on: on the one hand, that the BVET sample, in relation to these variables, is markedly homogeneous or that its variability does not respond to a pattern linked to the completion, repetition or abandonment of studies. This could be affected by the circumstance that these students begin the BVET at the age of compulsory school attendance and have been referred to this training after a pathway of failure in CSE. Furthermore, given the weakness in the significance of the two aforementioned variables, we must also accept that we are still facing an analysis model that needs other variables to better understand why these students finish, repeat or drop out. Future studies should delve into other aspects that may have a more significant influence, such as the possibility or not of finding employment through

informal or family networks, as well as the rural or urban context in which they live, and the disparity of employment opportunities offered to them.

In contrast, the IVET subsample seems to respond better to the full sample picture, which is not surprising given that they are overrepresented in the total sample. However, it presents some differences: in this group, the weight of the relationship with the teacher loses intensity, whereas the discriminatory capacity of the relationship with peers increases instead, which would indicate the fundamental role of this relationship in the affective bond that is established (González & Bernárdez, 2019) and that has the effect of a higher rate of completion of studies in those students who feel supported by their peers. On the other hand, expectations, future aspirations and achievements do not seem to show discriminatory capacity in this group, which would seem to indicate that these students show a greater intrinsic interest in the studies they receive, although we must admit that the measurement of student motivation does not reach any degree of significance when distinguishing between those who complete, repeat or drop out of the IVET.

As a final comment, the results obtained lead us to ask ourselves: What makes one person see in his or her training process the scaffolding for his or her professional future and another not? On what do each anchor their expectations? Our results seem to indicate that a key aspect, although not the only one, is the fact of endowing the training process with future expectations. Along these lines, Tarabini and Curran (2015, p. 21) remind us that:

although there are pragmatic, instrumental and rational elements in young people's decisions, these cannot be understood apart from the horizons of action that restrict and/or make them possible. Horizons that are the result of the intersection between the objective opportunities available to social actors and their subjective perceptions.

Therefore, in view of the dominant discourse that advocates increasing the educational level of the population, factors such as engagement, especially in its emotional and cognitive dimensions, must be taken into account in the design of educational policies, not only in terms of practices in the centres, but also in the territorial configuration of the offer, so that the educational experience has a sense of professional future for the students.

Funding

This article is based on the research project “Itinerarios de éxito y abandono en la Formación Profesional del sistema educativo de nivel 1 y 2” (Itineraries of success and dropout in Vocational Training in level 1 and 2 of the education system) (EDU2013-42854- R), funded under the R&D&I Programme oriented to the challenges of society of the Ministry of Economy, Industry and Competitiveness, State Research Agency and the European Regional Development Fund, which had continuity through the regional research project “Itinerarios de éxito y abandono en Formación Profesional de nivel 1 y 2 del sistema educativo de la provincia de Valencia” (Itineraries of success and dropout in Vocational Training in level 1 and 2 of the education system in the province of Valencia) (GV/2018/038), funded by the Ministry of Education, Research, Culture and Sport of the Valencia Regional Government.

Bibliographical references

Abiétar, M., & Navas, A. A. (2017). El sentido de la escolaridad obligatoria como transición o como fin. *Profesorado: revista de currículum y formación del profesorado*, 21(4), 75-94. <https://doi.org/10.30827/profesorado.v21i4.10046>

Bernad, J.C., & Molpeceres, M.A. (2006). Discursos emergentes sobre la educación en los márgenes del sistema educativo. *Revista de Educación*, 341, 149-169.

Bourdieu, P. (2001, 2^a Ed.). *Poder, derecho y clases sociales*. Desclée de Brouwer.

Casal, J., García, M., Merino, R., & Quesada, M. (2006). Aportaciones teóricas y metodológicas a la sociología de la juventud desde la perspectiva de la transición. *Papers* (79), 21-48. <https://doi.org/10.5565/rev/papers/v79n0.798>

Casal, J., Merino, R., & García, M. (2011). Pasado y futuro del estudio sobre la transición de los jóvenes. *Papers*, 96 (4), 1139-1162. <https://doi.org/10.5565/rev/papers/v96n4.167>

Cerdà-Navarro, A., Salva-Mut, F., & Comas Forgas, R. (2019). A typology of students in intermediate vocational education and training programmes based on student engagement factors, sociodemographic

characteristics, and intentions of dropping out. *European Journal of Education. Research, development and policy*, 54(4), 635-650. <https://doi.org/10.1111/ejed.12361>

De Puelles, M. (1999). *Educación e Ideología en la España contemporánea*. Tecnos.

Garcia, M., Casal, J., Merino, R., & Sánchez, A. (2013). Itinerarios de abandono escolar y transiciones tras la Educación Secundaria Obligatoria. *Revista de Educación*, 361, 65-94. <https://doi.org/10.4438/1988-592X-RE-2011-361-135>

Garcia, M., Sánchez-Gelabert, A., & Valls, O. (2022). Itinerarios educativos, trazabilidad y autopercepción de notas en las transiciones postobligatorias. *Revista de Educación* (396), 97-126. <https://doi.org/10.4438/1988-592X-RE-2022-396-531>

Garcia, M., & Valls, O. (2018). Trayectorias de permanencia y abandono educativo temprano: Análisis de secuencias y efectos de la crisis económica. *Metamorfosis. Revista del Centro Reina Sofía Sobre Adlescencia y Juventud*, (8), 129-143.

González, M^aT., & Bernárdez-Gómez, A. (2019). Elementos y aspectos del centro escolar y su relación con la desafección de los estudiantes. *Revista de Investigación en Educación*, 17(1), 5-19.

Instituto Nacional de Estadística (2019). *Encuesta de transición educativa-formativa e inserción laboral Año 2019. Nota de prensa*. https://www.ine.es/prensa/etefil_2019.pdf

Jacovkis, J., Montes, A., & Manzano, M. (2020). Imaginando futuros distintos. Los efectos de la desigualdad sobre las transiciones hacia la educación secundaria posobligatoria en la ciudad de Barcelona. *Papers*, 105(2), 279-302. <https://doi.org/10.5565/rev/papers.2773>

Ley Orgánica 3/2022 de ordenación e integración de la Formación Profesional. 1 de abril de 2022. Boletín Oficial del Estado No. 78.

Martínez, I., Bernad, J.C., Molpeceres, M.A., Abiétar, M., Navas, A., Marhuenda, F., & Giménez, E. (2015). Comprehensive education boundaries and remedies on the edges of the Spanish educational system. *European educational Research Journal*, 14 (3-4), 293-311. <https://doi.org/10.1177/1474904115590043>

Martínez, J. S., & Merino, R. (2011) Formación Profesional y desigualdad de oportunidades por clase social y género. *Támpora: revista de historia y sociología de la educación*, (14), 13-37.

Merino, R. (2013). Las sucesivas reformas de la formación profesional en España o la paradoja entre integración y segregación escolar. *Archivos Analíticos de Políticas Educativas*, 21(66). <https://doi.org/10.14507/epaa.v21n66.2013>

Merino, R., García, M., Casal, J. & Sánchez, A. (2011). Itinerarios formativos y laborales de los jóvenes graduados en formación profesional. sobre algunos prejuicios en la formación profesional. *Sociología del Trabajo*, (72), 137-156.

Ministerio de Educación y Formación Profesional (2021). *Sistema estatal de indicadores de la educación 2021*. <https://www.educacionyfp.gob.es/dam/jcr:98edb864-c713-4d48-a842-f87464dc8aee/seie-2021.pdf>

Navas, A. A., Abiéitar, M., Bernad, J. C., Córdoba, A. I., Giménez, E., Meri, E. & Quintana, E. (2021). Implicación del estudiantado en Formación Profesional: análisis diferencial en la provincia de Valencia. *Revista de Educación*, 394, 189-213. <https://doi.org/10.4438/1988-592X-RE-2021-394-505>

Niemeyer, B., & Colley, H. (2015). Why do we need (another) special issue on gender and VET? *Journal of Vocational Education & Training*, 67(1), 1-10. <https://doi.org/10.1080/13636820.2014.971498>

Nylund, M., Rosvall, P. Å., Eiríksdóttir, E., Holm, A. S., Isopahkala-Bouret, U., Niemi, A. M. & Ragnarsdóttir, G. (2018). The academic-vocational divide in three Nordic countries: Implications for social class and gender. *Education Inquiry*, 9(1), 97-121. <https://doi.org/10.1080/20004508.2018.1424490>

Reschly, A. L., & Christenson, S. L. (2012). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the *Engagement* Construct. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 3-20). Springer.

Tarabini, A., & Curran, M. (2015). El efecto de la clase social en las decisiones educativas: un análisis de las oportunidades, creencias y deseos educativos de los jóvenes. *Revista de Investigación en Educación*, 13(1), 7-26.

Tarabini, A., Curran, M., Montes, A., & Parcerisa, Ll. (2019). Can educational engagement prevent Early School Leaving? Unpacking the school's effect on educational success. *Educational Studies*, 45(2), 226-241. <https://doi.org/10.1080/03055698.2018.1446327>

Torrents, D., Merino, R., Garcia, M., & Valls, O. (2018). El peso del origen social y del centro escolar en la desigualdad de resultados al final de la escuela obligatoria. *Papers*, 103(1), 29-50. <https://doi.org/10.5565/rev/papers.2300>

UNESCO (2015). *Educación 2030. Declaración de Incheon y Marco de Acción para la realización del Objetivo de Desarrollo Sostenible 4*. https://unesdoc.unesco.org/ark:/48223/pf0000245656_spa

Contact address: Míriam Abiétar López, Universitat de València, Facultat de Filosofia i Ciències de l'Educació, Departament Didàctica i Organització Escolar. Av. Blasco Ibáñez, 30, CP 46010, Valencia, Spain. E-mail: miriam.abietar@uv.es

Diet quality and academic performance in schoolchildren: the moderating role of weight status

Calidad de la dieta y rendimiento académico en escolares: el papel moderador del estado de peso

<https://doi.org/10.4438/1988-592X-RE-2023-401-588>

Pedro José Carrillo-López

<https://orcid.org/0000-0003-0063-7645>

Government of the Canary Islands

Abstract

Background. There is currently a trend in the scientific literature that studies the association between dietary patterns and their impact on academic performance. **Aim.** To determine the predictive power of diet quality on academic performance in schoolchildren by observing the moderating role of weight status. **Methods.** A total of 244 schoolchildren (43.6% boys and 53.7% girls) from Spain (Tenerife), aged between 10 and 12 years ($M \pm SD: 10.80 \pm 1.81$ years) participated in this empirical descriptive and cross-sectional ex post facto study. Academic achievement was calculated through the grade obtained in the subjects of the primary school curriculum (Royal Decree 126/2014), the *KIDMED* questionnaire was used to quantify diet quality and nutritional status was assessed through the *Body Mass Index* (kg/m^2) adjusted for sex and age. **Results.** Higher diet quality correlates with higher scores in all academic subjects ($p < .05$) with the exception of Natural Sciences, Art Education, and Religion/Values and French ($p > .05$) with and without adjusting the model for normal weight. After adjusting the model for overweight, the significant relationship disappears in all subjects ($p > .05$), except for Mathematics and English ($p < .05$, for both). Likewise, the Johnson-Neyman test showed that there is no moderation of weight status in relation to diet quality and academic subjects ($p > .05$); with the exception of Physical Education ($p < .05$). **Conclusions.** Based on these results, further intervention and longitudinal research should be conducted to reveal possible strategies and policies that would improve school performance and overall health across the lifespan.

Keywords: Mediterranean diet, obesity, academic performance, nutritional education, childhood.

Resumen

Antecedentes. Actualmente existe una tendencia en la literatura científica que estudia la asociación entre los patrones dietéticos y su impacto en el rendimiento académico. Objetivo. Determinar el poder predictivo de la calidad de la dieta sobre el rendimiento académico en escolares observando el papel moderador del estado ponderal. Métodos. Un total de 244 escolares (43.6% niños y 53.7% niñas) de España (Tenerife), con edades comprendidas entre los 10 y los 12 años ($M \pm DE: 10.80 \pm 1.81$ años) participaron en este estudio empírico descriptivo y transversal *ex post facto*. El rendimiento académico se calculó a través de la calificación obtenida en las asignaturas del currículo de primaria (Real Decreto 126/2014), se utilizó el cuestionario KIDMED para cuantificar la calidad de la dieta y se valoró el estado nutricional a través del Índice de Masa Corporal (kg/m^2) ajustado por sexo y edad. Resultados. Una mayor calidad de la dieta se correlaciona con mayores puntuaciones en todas las asignaturas académicas ($p < .05$) a excepción de Ciencias Naturales, Educación Artística y Religión/Valores y Francés ($p > .05$) con y sin ajustar el modelo por peso normal. Tras ajustar el modelo por sobrepeso, la relación significativa desaparece en todas las asignaturas ($p > .05$), excepto en Matemáticas e Inglés ($p < .05$, para ambas). Asimismo, la prueba de Johnson-Neyman mostró que no existe moderación del estado de peso en relación con la calidad de la dieta y las asignaturas académicas ($p > .05$); a excepción de Educación Física ($p < .05$). Conclusiones. En base a estos resultados, se deben realizar más intervenciones e investigaciones longitudinales para revelar posibles estrategias y políticas que mejoren el rendimiento escolar y la salud general a lo largo de la vida.

Palabras clave: Dieta mediterránea, obesidad, rendimiento académico, educación nutricional, infancia.

Introduction

From the heart to the earth through the path of culture, the Mediterranean diet is a cultural heritage recognised by UNESCO that looks towards a healthy future (Serra-Majem & Ortiz-Andrellucchi, 2018). However, the dietary profile of the school population is increasingly far from the quality of the optimal diet, mainly due to the low consumption of fruit,

vegetables, whole grains and the high consumption of meat and meat products, sweets and sugary drinks (García-Cantó et al., 2019). This variability in the quality of healthy and/or sustainable diets varies widely across the world and contexts (Carrillo-López et al., 2021a). In all of them, it has been described that this low diet quality can affect the survival, growth and proper development of schoolchildren by being associated with increased blood pressure and metabolic risk (Teixeira et al., 2021). In other words, energy-rich, nutrient-poor, ultra-processed foods are causing a global epidemic of disease (Kupka et al., 2020). Therefore, improving dietary habits towards optimal diet quality could be associated with better health from a holistic view (García-Hermoso et al., 2022).

Specifically, at the mental level, Kaliszewska et al. (2021) reveal a strong link between nutrition, mitochondrial functioning and cognition. In this sense, the integration of a healthy diet can provide optimal conditions for brain development and teach (Naveed et al., 2020). In this regard, there is currently an important scientific stream studying the association between dietary patterns and school academic performance. This literature has mainly focused on adolescents; presenting contradictory information on this relationship. That is, in 2016, a Europe-wide study pointed out that adherence to the Mediterranean diet may have a beneficial influence on academic performance (in four indicators: mathematics, language, mathematics and language average and grade point average). They highlighted that the benefits of adherence to the Mediterranean diet on academic performance may be stronger as young people adhere to optimal levels of the Mediterranean diet (Esteban-Cornejo et al., 2016).

A 2017 systematic review showed moderate associations for dietary intakes characterised by regular breakfast consumption, lower intake of energy-rich and nutrient-poor foods, and overall diet quality with respect to academic performance outcomes (Burrows et al., 2017). A three-year longitudinal study of high school students found that dietary habits were correlated with academic performance only in boys (Dubuc et al., 2020). Other research with high school students and standardised academic aptitude scores (Nawabjan & Nazni, 2021; Al-Saadi et al., 2020; Kristo et al., 2020) or global ratings of academic performance showed positive relationships with diet quality (Maniaci et al., 2021). In contrast, high academic performance has been found to be associated with a higher incidence rate of food addiction detection (Tserne et al., 2021).

Fewer studies have been found that analyse this relationship in primary school children. Among them, Qasrawi et al. (2021) found that healthy nutrition significantly predicts achievement scores. In both boys and girls, high academic performance was associated with high consumption of fruits and vegetables, low consumption of soft drinks, beverages (juices with sugar) and energy drinks, respectively. The high diet quality group scored higher in Arabic, English, mathematics, science and mean total score. Similarly, Faught et al. (2017) reported that poor diet quality was less likely to result in lower overall school performance. Similarly, optimal diet quality, such as not consuming cakes and sweet pastries, Chips, fried dumplings and pastries, was associated with differences in academic performance but only in mathematics (Gaete-Rivas et al., 2021). In contrast, another study found that a healthier diet was associated with better reading skills, but not with arithmetic skills (Haapala et al., 2017). However, another study found no significant relationship (Carrillo-López et al., 2021b). Similarly, Nakahara et al., (2020) did not observe a significant association between candy consumption and mathematics scores, however, higher candy consumption was significantly associated with higher Mongolian language scores.

Some studies have found that this association may be moderated by other factors. For example, in several studies after adjusting for gender, age or place of residence, the association was not significant (Mazandaranian et al., 2021; Barrios & Vernetta, 2021). Another possible moderator studied in secondary education has been weight status since it is obtained from the energy balance of nutrients ingested in the diet minus caloric expenditure and is measured by indicators such as body mass index for age and sex, an aspect that allows the diagnosis of weight status (Tapia-Serrano et al., 2021). In this study, they showed that language, mathematics and English scores were significantly associated with diet quality, independent of weight status. Despite the lack of significance, this study suggests that it seems likely that weight status might moderate the relationship between diet quality and academic performance. For this reason, they suggest further research and in primary school children. In this regard, this study is the first research to test the moderating role of weight status in the association between diet quality and academic performance in primary school children, and to examine how the relationship between diet quality and academic performance varies by weight status (normal weight vs. overweight). The choice of schoolchildren at this pre-adolescent stage is because this period is critical for neurodevelopment,

which is characterised by the establishment of behavioural patterns that can affect the physical, mental and academic health of schoolchildren, both immediately and later in life (Meli et al., 2022).

Based on these precedents, the predictive power of diet quality on academic performance in all subjects taken by primary school children is further explored by looking at the moderating role of weight status.

Method

Type of study and participants

Prior to conducting this research, the sample size was calculated in order to ensure robust results (Quispe et al., 2020). After jointly estimating the u (in reference to the number of variables) and f^2 (effect size in linear regression models) statistics, it was obtained that the minimum sample had to be a total of 217 subjects in order to carry out the linear regression technique, something that is fulfilled since we have a total sample of 244 students.

In this regard, a total of 244 schoolchildren (43.6% boys and 53.7% girls) from Spain (Tenerife), aged between 10 and 12 years ($M \pm SD: 10.80 \pm 1.81$ years) participated in this empirical descriptive and cross-sectional *ex post facto* study. Sampling was non-probabilistic, non-random and convenience sampling (access to the sample). Four public schools were selected from rural (two schools) and urban (two schools) environments. These schools have a medium socio-economic level. In previous meetings held with the school principals and legal guardians of the schoolchildren, they were informed of the study protocol and informed consent was requested so that the schoolchildren could participate. Inclusion criteria were considered to be between 10-12 years of age and not to suffer from previous pathologies. In addition, the following exclusion criteria were considered: I) Failure to present informed consent to participate in the research, II) Attendance at more than 90% of academic classes.

Procedure and instruments

The work was carried out during the months of March and April of the 2019/2020 academic year. Each head of school and the representatives

of the parents' associations were informed of the purpose and protocol of the research at a meeting. The working team consisted of a principal researcher and four collaborating colleagues (teachers specialising in Primary Education and Physical Education). A theoretical session was held prior to the completion of the KIDMED questionnaire with each study group in order to ensure that all participants understood the questionnaires in this study. The research team administered the test in the natural groups of the PE class. All questionnaires were administered during the first school session in order to avoid the possible fatigue of the school day and to interrupt the school dynamics as little as possible.

The research was carried out in accordance with the ethical standards recognised by the Declaration of Helsinki (2013 revision), following the recommendations of Good Clinical Practice of the EEC (document 111/3976/88 of July 1990) and the current Spanish legal regulations governing clinical research on humans (Royal Decree 561/1993 on clinical trials).

Academic performance was considered as the dependent variable in this study. This was assessed by means of the grade obtained by fifth and sixth grade students in the first and second assessment carried out in the subjects of the Primary Education curriculum: Natural Sciences, Social Sciences, Spanish Language and Literature, Mathematics, First Foreign Language: English, Physical Education, Religion/Values and Second Foreign Language: French (Royal Decree 126/2014, of 28 February, establishing the basic curriculum of Primary Education and Decree 89/2014, of 1 August, establishing the organisation and curriculum of Primary Education in the Autonomous Community of the Canary Islands).

The values of all variables ranged from one to ten points. The score obtained in each subject derives from the total assessment of the learning established in the assessment criteria described in the Resolution of 13 May 2015 establishing the rubrics of the assessment criteria in Primary Education in the Autonomous Community of the Canary Islands, which emanate directly from Royal Decree 126/2014, of 28 February. These criteria are the fundamental reference for the assessment and marking of pupils throughout Spain.

The independent variables are:

I) Diet quality, quantified using the KIDMED questionnaire (Serra et al., 2004). This instrument is composed of 16 items that represent standards

of the traditional Mediterranean diet. Four of them are assessed with a negative score (-1 point) if answered positively (items 6, 12, 14 and 16), while the remaining twelve items are assessed with a positive score (+ 1) if answered positively. After summation, an overall score between -4 and 12 is obtained, which describes a better or worse quality of the diet. The value of the KIDMED index is: score ≤ 3 indicating a very low quality diet; score between 4 and 7 indicating the need to improve the dietary pattern to conform to the Mediterranean model; and finally, score ≥ 8 , showing an optimal Mediterranean diet. In order to obtain greater statistical power, participants were categorized into two groups: improvable DQ (≤ 7) and optimal DQ (≥ 8) as has been done in other previous studies (Carrillo-López, 2023; Tapia-Serrano et al., 2021).

II) Weight and height were determined using an electronic scale (TANITA TBF 300A, USA) and measuring rod (SECAA800, USA) with an accuracy of 100 g and 1 mm respectively, following the protocol of the International Society for the Advancement of Kynanthropometry with level I certified personnel. From these anthropometric variables, the body mass index (kg/m^2) was calculated. From this index, age- and sex-adjusted nutritional status was diagnosed (Cole & Lobstein, 2012). Participants were categorised into two groups: normal weight and overweight (overweight + obesity).

Statistical analysis

The normality and homogeneity of variances were obtained through the Kolmogorov Smirnov ($p > .212$) and Levene statistics ($p > .774$), respectively. As a normal distribution of the values recorded was observed, a parametric analysis was chosen. A differential analysis was then carried out. To indicate the characteristics of this sample, frequency distribution was used for categorical variables and descriptive analyses using the mean \pm standard deviation for continuous variables. Student's t-test was used to test for significant differences between groups and the chi-squared test for categorical variables. Subsequently, the PROCESS macro tool (version 3.5) of the SPSS software (IBM Corp, Armonk, New York, USA (version 23) was applied. This tool allows for moderation analysis (Hayes, 2017).

The initial analysis indicated no significant differences between the sexes; consequently, all analyses were conducted with males and females

together. This is to obtain greater statistical power. The moderation analysis was used to analyse whether diet quality (independent variable) was linked to academic performance (dependent variable) by looking at the moderating effect of weight status (moderator variable). Before interpreting the coefficients, goodness-of-fit and model assumptions were assessed. To analyse the goodness-of-fit, the F-test was used, which indicates whether the linear relationship being analysed is statistically significant. It should be noted that this statistic was significant, thus confirming the relevance of the regression technique (Martínez et al., 2020). With respect to the assumptions, as indicated by Pardo & San Martín (2010), the assumption of non-collinearity, linearity, independence of the errors and the Breusch-Pagan test to check the homogeneity of the residuals were checked, as well as the Breusch-Pagan test to check the homogeneity of the residuals. These assumptions are met in all models. The influence of outliers was also tested using Cook's distance. As the value obtained was less than 1, it was concluded that there was no influential case. An ordinary least squares (OLS) regression analysis was performed to predict the continuous variables (academic performance and KID-MED). In turn, since it is essential to determine the contribution of the predictor when carrying out a regression model, the R^2 statistic (Raschka & Mirjalili, 2019) was used (Raschka & Mirjalili, 2019). SPSS (version 23) was used for the analysis of all data. However, for the calculation of the sample size, the programme R, version 4.1.2 (pwr package) was used (Champely et al., 2018) with the significance level set at 5% ($p \leq .05$).

Results

Table I shows the scores obtained in the different variables of the study according to sex. It should be noted that no significant differences were obtained for any variable according to sex ($p > .05$).

In order to determine the predictive power of diet quality on academic performance in the core subjects by looking at the moderating role of weight status, a linear regression test was applied (see Table II). The crude model showed that higher values for diet quality correlated with higher values for academic performance in all subjects ($p < .05$) except Natural Science ($p > .05$). However, after the model was adjusted for Overweight, this relationship between diet quality and academic

TABLE I. Scores obtained in the different variables of the study according to sex

	Males M ± SD (n = 118)	Females M ± SD (n = 126)	F	p	d
Natural Sciences (1-10) ^a	5.69 ± 6.88	1.70 ± 1.40	4.132	.094	0.15
Social Sciences (1-10) ^a	6.07 ± 1.50	7.05 ± 1.57	1.050	.823	0.07
Spanish Language and Literature (1-10) ^a	5.76 ± 1.85	6.79 ± 1.69	1.793	.183	0.11
Mathematics (1-10) ^a	5.73 ± 1.99	6.48 ± 1.72	1.457	.230	0.10
English (1-10) ^a	5.61 ± 1.52	6.79 ± 1.45	1.591	.443	0.09
Art Education (1-10) ^a	5.87 ± 1.18	6.74 ± .97	1.928	.168	0.10
Physical Education (1-10) ^a	6.85 ± 1.30	7.09 ± 1.20	1.289	.592	0.07
Religion/ Values (1-10) ^a	6.30 ± 1.49	6.77 ± 1.57	1.041	.947	0.05
French (1-10) ^a	5.79 ± 1.28	6.90 ± 1.05	4.481	.064	0.15
Age (years)	10.44 ± 0.71	10.29 ± 0.84	1.619	.365	0.10
Height (cm)	154.01 ± 8.41	153.51 ± 8.28	1.570	.758	0.09
Weight (kg)	52.19 ± 13.25	48.11 ± 12.53	1.036	.218	0.12
BMI (kg/m ²) ^c	21.84 ± 4.22	22.28 ± 4.09	1.968	.044	0.16
Normal weight (n = 124) ^d	25%	25.9%	-	.389	-
Overweight ^d (n = 120)	27.6%	21.6%	-	.159	-
DQ ^e	7.53 ± 2.07	7.84 ± 2.62	2.935	.877	0.07
Improveable (n=108) ^d	25%	21.6%	-	.584	-
Optimal (n=136) ^d	27.6%	25.9%	-	.327	-

M ± SD = mean ± standard deviation. ^a Mean score obtained in the subject. ^c BMI = Body Mass Index. ^d Data presented as a cumulative percentage of the total sample. ^e Diet quality expressed from the mean score of the KIDMED scale.
Source: Compiled by author.

performance disappeared for all subjects ($p > .05$) except Mathematics and English ($p < .05$).

On the other hand, after observing the predictive power of diet quality on academic performance in the specific subjects of the curriculum by observing the moderating role of weight status (see Table III), significant values were obtained in the raw model only for the area of Physical Education ($p < .05$), but this relationship disappeared when the model was adjusted to overweight ($p > .05$).

TABLE II. Predictive value of diet quality on academic performance in core subjects

	Natural Sciences	Social Sciences	Spanish Language and Literature	Mathematics	English
Model I ^a	$B = 0.084$	$B = 0.125$	$B = 0.173$	$B = 0.136$	$B = 0.140$
	$t = 1.258$	$t = 1.979$	$t = 2.392$	$t = 1.790$	$t = 3.439$
	$R^2 = .014$	$R^2 = .033$	$R^2 = .048$	$R^2 = .072$	$R^2 = .041$
	$p = .211$	$p = .041^*$	$p = .018^*$	$p = .046^*$	$p = .028^*$
Model II ^b	$B = 0.089$	$B = 0.152$	$B = 0.194$	$B = 0.142$	$B = 0.142$
	$t = 1.411$	$t = 2.120$	$t = 2.451$	$t = 1.814$	$t = 3.514$
	$R^2 = .016$	$R^2 = .028$	$R^2 = .051$	$R^2 = .078$	$R^2 = .051$
	$p = .192$	$p = .041^*$	$p = .014^*$	$p = .045^*$	$p = .024^*$
Model III ^c	$B = 0.044$	$B = 0.108$	$B = 0.148$	$B = 0.173$	$B = 0.173$
	$t = .439$	$t = 1.104$	$t = 1.311$	$t = 1.396$	$t = 1.622$
	$R^2 = .003$	$R^2 = .022$	$R^2 = .030$	$R^2 = .052$	$R^2 = .056$
	$p = .663$	$p = .274$	$p = .195$	$p = .048^*$	$p = .032^*$

Note. ^a Model I adjusted for age, gender and place of residence; ^b Model II adjusted for age, gender, place of residence and normal weight. ^c Model III adjusted for age, gender, place of residence and overweight.

Source: Compiled by author.

TABLE III. Predictive value of diet quality on academic performance in specific subjects

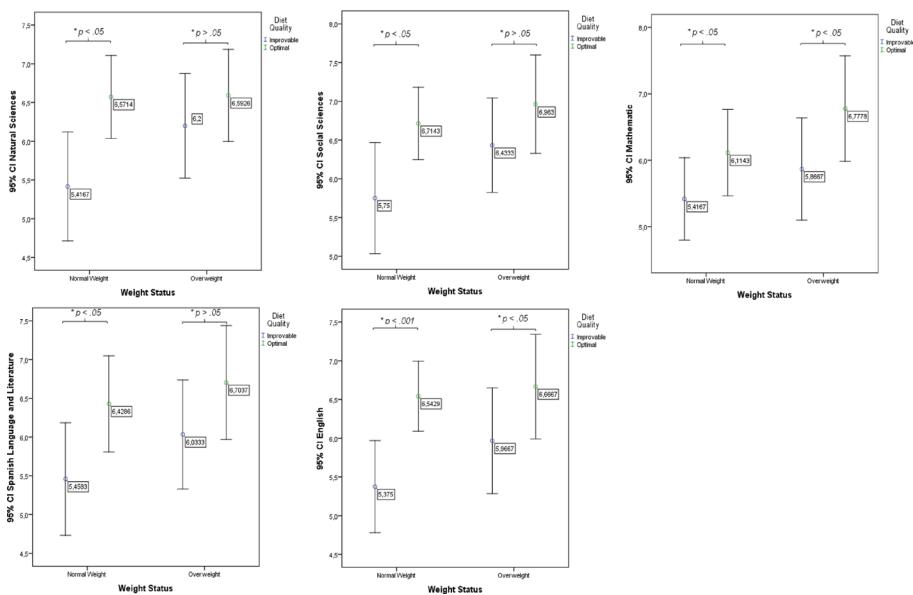
	Art Education	Physical Education	Religion/Values	French
Model I ^a	$B = 0.059$	$B = 0.127$	$B = 0.114$	$B = 0.037$
	$t = 1.749$	$t = 2.584$	$t = 1.877$	$t = 1.706$
	$R^2 = .014$	$R^2 = .055$	$R^2 = .030$	$R^2 = .004$
	$p = .206$	$p = .011^*$	$p = .063$	$p = .482$
Model II ^b	$B = 0.62$	$B = 0.124$	$B = 0.154$	$B = 0.123$
	$t = 1.825$	$t = 2.874$	$t = 1.982$	$t = 1.888$
	$R^2 = .014$	$R^2 = .062$	$R^2 = .032$	$R^2 = .005$
	$p = .192$	$p = .004^*$	$p = .056$	$p = .328$
Model III ^c	$B = 0.92$	$B = 0.111$	$B = 0.064$	$B = 0.038$
	$t = 1.221$	$t = 1.448$	$t = 1.645$	$t = 1.850$
	$R^2 = .026$	$R^2 = .037$	$R^2 = .008$	$R^2 = .003$
	$p = .227$	$p = .153$	$p = .522$	$p = .562$

Note. ^a Model I adjusted for age, gender and place of residence; ^b Model II adjusted for age, gender, place of residence and normal weight. ^c Model III adjusted for age, gender, place of residence and overweight.

Source: Compiled by author.

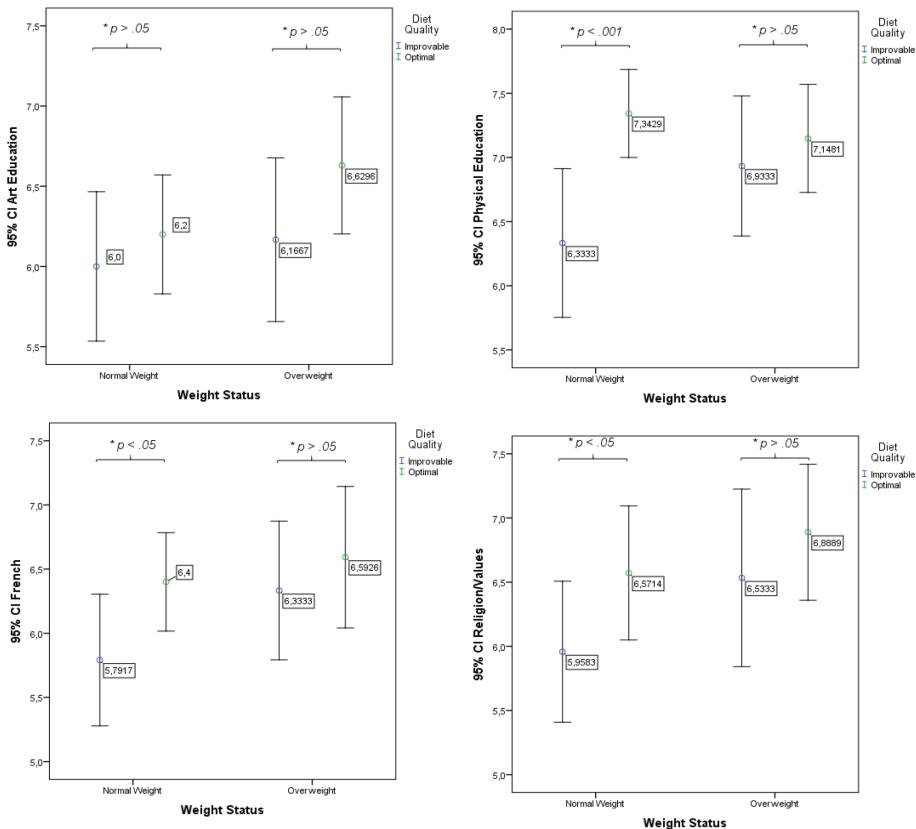
Figures I and II show the differences in academic performance between the diet quality groups (*improvable vs. optimal*) according to weight status (*normal-weight and overweight*). Normal-weight schoolchildren with good diet quality showed significantly higher scores in all academic scores compared to those with an improvable diet quality (with the exception of Art Education). However, those who were overweight showed no significant differences in academic subject scores (with the exception of Mathematics and English). It is prescriptive to note that interactions between weight status and diet quality (the Johnson–Neyman method was performed; X^*W) were calculated in relation to academic subjects and were not significant for any subject (all $p > .05$) with the exception of Physical Education [$\beta = .124$; 95% CI = .102, .142].

FIGURE I. Differences in core subjects between adherence to the Mediterranean diet by weight status



Source: Compiled by author

FIGURE II. Differences in specific subjects between adherence to the Mediterranean diet groups by weight stat



Source: Compiled by author

Discussion

The aim of this study was to determine the predictive power of diet quality on academic performance in all subjects taken by primary school children by looking at the moderating role of weight status. The main findings reveal that higher diet quality correlates with higher scores in all academic subjects with the exception of Natural Sciences, Art Education, Religion/Values and French.

As can be seen, there is a stronger relationship of diet quality with subjects that have a higher curricular weight. These results can be explained through neuroscience, understood as a discipline that combines psychology, pedagogy and neuroscience to explain how the brain works in learning processes. In this sense, our brain tends to understand better what is perceived by the senses, i.e. the concrete as opposed to the abstract. In this respect, exact areas such as mathematics or language study on many occasions' abstract properties, structures and relationships that require a higher level of difficulty than experimental sciences such as natural sciences, which resort to experiments or controlled tests to make a discovery (Decree 89/2014, of 1 August). In this line of argument, these results may be due to the fact that a higher quality of diet is positively related to elaboration strategies, organisational strategies, critical thinking, self-regulation, time and study habits, self-regulation of effort, and intrinsically oriented objectives (Chacón-Cuberos et al., 2018).

These aspects coincide with the findings of a meta-analysis, which indicates that improved dietary habits could be associated with more physically and mentally active behaviours, and thus lead to better overall health; aspects that could undoubtedly have an impact on higher academic performance. On the other hand, at the psychological level, Ekman (2021) indicates that thoughts, emotional patterns and psychological dynamics are strongly interrelated with learning. That is, if a person is aware of how they can influence and regulate their situation, they gain a sense of control and self-awareness that can be extrapolated to other contexts of their life, such as academic performance.

At the physiological level, a meta-analysis indicated that various micronutrients such as vitamin B12, zinc or iron may have a positive effect on subjects such as Mathematics, English, Geography, Science and Arts by improving certain executive functions (Meli et al., 2022). Other studies have also examined the relationship between breakfast and snack consumption and academic and cognitive performance in adolescent students, finding positive associations (Masoomi et al., 2020; Gaylor et al., 2021). Specifically, healthy eating behaviours predicted a greater increase in task-related activation of the right dorsolateral prefrontal cortex (Papasideris et al., 2020).

In turn, dietary patterns with low intakes of fish, fruit and vegetables, and high in fast food, sausages and soft drinks have been linked to poor cognition and academic performance. These differences remained significant for both the normal-weight and overweight groups. Similarly, break-

fast intake was associated for mathematics and science achievement (Vik et al., 2022). This study reported in mediation analyses that being hungry at school explained one-third of the decline in science achievement and more than half of the decline in mathematics achievement from 2015 to 2019. This could explain why in the present study, after adjusting the model for overweight, the significant relationship disappears in all subjects except Mathematics and English. Likewise, looking at the moderating role of weight status, the Johnson-Neyman method showed that there is no moderation of weight status in relation to diet quality and academic subjects, with the exception of Physical Education.

These results are partially similar to those obtained by Tapia-Serrano et al., (2021) where no interaction was found between weight status and diet quality in relation to academic indicators. This result may be due to the fact that the predominant methodology in the area of Physical Education is motor play and the attention for understanding the game and playing is high in all students (Rosa et al., 2018). That is to say, through play, the activation of perception and decision-making mechanisms is made possible, as well as the development of execution capacities; the acquisition of new motor skills is accessed, concepts are contrasted and the motivation of pupils to participate in activities and tasks that have an eminently recreational character, typical of the use of play and the forms played, without losing their priority function of training, is activated. In this sense, it has been described that the quality of the diet makes it possible to have more energy to carry out motor games and enables a greater capacity for attention. Therefore, it may have robust power to predict later academic success (Rosa-Guillamón et al., 2020). Hence, as schoolchildren have the necessary energy to play in Physical Education, there may be a relationship between diet quality and academic performance in this area.

In this study, in the predictive and differential analysis, it was observed that after adjusting the model to normal weight, this relationship between diet quality and academic performance is intensified in all academic subjects with the exception of Natural Sciences, Art Education, Religion/Values and French.

In this regard, Gabbianelli & Damiani (2018) suggest that the link between nutrition and learning may derive from the close interconnection between gut microbiota and cognitive mechanisms transiting the gut-brain axis. This study in adults provides evidence that 5-HT and BDNF mediate the association between overweight/obesity and executive control.

Therefore, these biological pathways may provide valid scientific data on the link between overweight/obesity and executive control (Si et al., 2021). However, Hernández-García et al., (2020) on the basis of the social structuring theory developed by Anthony Giddens indicate that academic performance and its relationship with health is a multifactorial problem, as it takes into account various factors: personal, social and institutional, as the individual and society work together.

In this sense, under the protection of the new educational law (Organic Law 3/2020, of December 29), the educational administrations must adopt measures so that physical activity and healthy eating are part of the behavior of children and young people during the school day, in the terms and conditions that, following the recommendations of the competent bodies, guarantee an adequate development to promote a healthy and autonomous life. Therefore, the educational system, in cooperation with the health system, must take responsibility for the physical and nutritional development of students insofar as they fall within their legal obligations. Undoubtedly, these professionals must assume to carry out intervention days in these early age stages in order to achieve the core objectives of the 2030 Agenda for Sustainable Development (Carrillo-López, 2022; García-Hermoso et al., 2022)".

Nevertheless, these findings should be interpreted with caution due to the fact that this study was not interventionist, but based on self-reported data, with unknown quality and quantity of food consumed daily by schoolchildren. In addition, the low sample size is undoubtedly another limitation. Similarly, it is difficult to infer a cause and effect relationship between attention and academic performance, since, as we have seen, there are confounding factors that are likely to influence these relationships and have not been considered in this study (such as socio-economic status). Thus, these effects could be related to environmental aspects and deserve to be further investigated in future studies.

Conclusion

Following the results obtained, it is concluded that: I) higher diet quality in primary school children is related to higher scores in all academic subjects with the exception of Natural Sciences, Art Education, Religion/Values and French; II) there is no moderation of weight status in relation to diet quality and academic subjects; III) students with normal weight

status who have good diet quality show significantly higher scores in all academic subjects compared to those with poor diet quality (except for Art Education). However, those who are overweight do not show significant differences in academic subject scores (except Mathematics and English) considering their diet quality. Based on these results, further intervention research should be conducted to reveal possible strategies and policies that would enhance positive behaviour change in relation to the recommended accurate dietary intake, with the aim of improving school performance and overall health throughout life. The awareness that habits during childhood become elements of daily life in adulthood, makes us realise the importance of the long-term consequences of maintaining healthy lifestyle habits, such as acceptable diet quality during childhood and adolescence on academic performance, hence the importance of these results.

Bibliographic references

Al-Saadi, L. S., Ali, A., Waly, M. I., & Al-Zuhaibi, K. M. (2020). Impact of dietary patterns and nutritional status on the academic performance of Omani school students. *Journal of Pharmacy and Nutrition Sciences*, 10(3), 74-87. <https://doi.org/10.29169/1927-5951.2020.10.03.1>

Barrios, E. M., & Vernetta, M. (2021). Adherencia a la dieta mediterránea y rendimiento académico en escolares de 1º y 2º de educación secundaria obligatoria. *Nutrición Clínica y Dietética Hospitalaria*, 41(3).

Burrows, T., Goldman, S., Pursey, K., & Lim, R. (2017). Is there an association between dietary intake and academic achievement: a systematic review. *Journal of Human Nutrition and Dietetics*, 30(2), 117-140. <https://doi.org/10.1111/jhn.12407>

Carrillo-López, P. J. (2023) Attention and academic performance: The moderator role of weight status and diet quality. *International Journal of Instruction*, 16(1), 707-722. <https://doi.org/10.29333/iji.2023.16139a>

Carrillo-López, P. J. (2022). Capacidad atencional, estado de peso y calidad de la dieta en escolares. *Apunts. Educación física y deportes*, 4(150), 1-9. [https://doi.org/10.5672/apunts.2014-0983.es.\(2022/4\).150.01](https://doi.org/10.5672/apunts.2014-0983.es.(2022/4).150.01)

Carrillo-López, P. J., Perujo, M., García-Cantó, E., & Rosa-Guillamón, A. (2021). Diferencias en la calidad de la dieta de escolares españoles de

educación primaria de entornos rurales y urbanos. *Revista Española de Nutrición Comunitaria*, 27(1), 8-14. <https://doi.org/10.33155/j.ramd.2021.06.002>

Carrillo-López, P. J., Rosa-Guillamón, A., Cantó, E., García, J. E., & Soto, J. J. (2021). Analysis of the quality of diet and academic performance in rural primary school students. *NeuroRegulation*, 8(1), 14-14. <https://doi.org/10.15540/nr.8.1.14>

Chacón-Cuberos, R., Zurita-Ortega, F., Martínez-Martínez, A., Olmedo-Moreno, E. M., & Castro-Sánchez, M. (2018). Adherence to the Mediterranean diet is related to healthy habits, learning processes, and academic achievement in adolescents: a cross-sectional study. *Nutrients*, 10(11), 1566. <https://doi.org/10.3390/nu10111566>

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., et al. (2018). Package ‘pwr.’ <http://cran.r-project.org/package=pwr>

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. LEA.

Cole, T.J., & Lobstein, T (2012). Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity*, 7(4), 284-294. <https://doi.org/10.1111/j.2047-6310.2012.00064.x>

Dubuc, M. M., Aubertin-Leheudre, M., & Karelis, A. D. (2020). Life-style habits predict academic performance in high school students: The adolescent student academic performance longitudinal study (ASAP). *International journal of environmental research and public health*, 17(1), 243. <https://doi.org/10.3390/ijerph17010243>

Ekman, R., Fletcher, A., Giota, J., Eriksson, A., Thomas, B., & Bååthe, F. (2021). A Flourishing Brain in the 21st Century: A Scoping Review of the Impact of Developing Good Habits for Mind, Brain, Well-Being, and Learning. *Mind, Brain, and Education*. <https://doi.org/10.1111/mbe.12305>

Esteban-Cornejo, I., Izquierdo-Gómez, R., Gómez-Martínez, S., Padilla-Moledo, C., Castro-Piñero, J., Marcos, A., & Veiga, O. L. (2016). Adherence to the Mediterranean diet and academic performance in youth: the UP&DOWN study. *European journal of nutrition*, 55(3), 1133-1140. <https://doi.org/10.1007/s00394-015-0927-9>

Faught, E. L., Williams, P. L., Willows, N. D., Asbridge, M., & Veugelers, P. J. (2017). The association between food insecurity and academic

achievement in Canadian school-aged children. *Public health nutrition*, 20(15), 2778-2785. <https://doi.org/10.1017/S1368980017001562>

Gabbianelli, R., & Damiani, E. (2018). Epigenetics and neurodegeneration: Role of early-life nutrition. *The Journal of nutritional biochemistry*, 57, 1-13. <https://doi.org/10.1016/j.jnutbio.2018.01.014>

Gaete-Rivas, D., Olea, M., Meléndez-Illanes, L., Granfeldt, G., Sáez, K., Zapata-Lamana, R., & Cigarroa, I. (2021). Hábitos alimentarios y rendimiento académico en escolares chilenos de quinto a octavo año básico. *Revista chilena de nutrición*, 48(1), 41-50. <https://doi.org/10.4067/S0717-75182021000100041>

García-Cantó, E., Carrillo-López, P. J., & Rosa-Guillamón, A. (2019). Análisis de la dieta mediterránea en escolares de Primaria, Secundaria y Bachillerato. *Revista chilena de nutrición*, 46(4), 469-476. <https://doi.org/10.4067/S0717-75182019000400469>

García-Hermoso, A., Ezzatvar, Y., López-Gil, J. F., Ramírez-Vélez, R., Olloquequi, J., & Izquierdo, M. (2022). Is adherence to the Mediterranean diet associated with healthy habits and physical fitness? A systematic review and meta-analysis including 565 421 youths. *British Journal of Nutrition*, 1-12. <https://doi.org/10.1017/S0007114520004894>

Gaylor, C., Young, H., & Benton, D. (2021). The Impact of Glycemic Load on Cognitive Performance in Adults: A Systematic Review and Meta-Analysis. *Current Developments in Nutrition*, 905-905. https://doi.org/10.1093/cdn/nzab049_018

Haapala, E. A., Eloranta, A. M., Venäläinen, T., Jalkanen, H., Poikkeus, A. M., Ahonen, T., et al. (2017). Diet quality and academic achievement: a prospective study among primary school children. *European journal of nutrition*, 56(7), 2299-2308. <https://doi.org/10.1007/s00394-016-1270-5>

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hernández-García, E., Álvarez-Martínez, B., & Ruiz-Ruisánchez, A. (2020). Health and performance in students of different academic levels. *Revista del Hospital Juárez de México*, 86(4), 196-201.

Kaliszewska, A., Allison, J., Martini, M., & Arias, N. (2021). The Interaction of Diet and Mitochondrial Dysfunction in Aging and Cognition. *International Journal of Molecular Sciences*, 22(7). <https://doi.org/10.3390/ijms22073574>

Kristo, A. S., Gültekin, B., Öztağ, M., & Sikalidis, A. K. (2020). The effect of eating habits' quality on scholastic performance in Turkish adolescents. *Behavioral sciences*, 10(1), 31. <https://doi.org/10.3390/bs10010031>

Kupka, R., Siekmans, K., & Beal, T. (2020). The diets of children: overview of available data for children and adolescents. *Global Food Security*, 27, 100442. <https://doi.org/10.1016/j.gfs.2020.100442>

Maniaci, G., La Cascia, C., Giannanco, A., Ferraro, L., Palummo, A., Saia, G. F., et al. (2021). The impact of healthy lifestyles on academic achievement among Italian adolescents. *Current Psychology*, 1-7. <https://doi.org/10.1007/s12144-021-01614-w>

Martínez, M., Sánchez-Villegas, Toledo, E., & Faulín, J. (2020). *Bioestadística amigable*. ELSEVIER

Masoomi, H., Taheri, M., Irandoust, K., H'Mida, C., & Chtourou, H. (2020). The relationship of breakfast and snack foods with cognitive and academic performance and physical activity levels of adolescent students. *Biological Rhythm Research*, 51(3), 481-488. <https://doi.org/10.1080/09291016.2019.1566994>

Mazandaranian, M. R., Omidvar, N., Eini-Zeinab, H., Doustmohammadian, A., Ashoori, M., Abdar-Esfahni, B., et al. (2021). Associations between Diet Quality, Weight Status and Academic Performance in Senior High-School Students in Tehran. *Nutrition and Food Sciences Research*, 8(2), 27-33. <https://doi.org/10.52547/nfsr.8.2.27>

Meli, A. M., Ali, A., Mhd Jalil, A. M., Mohd Yusof, H., & Tan, M. (2022). Effects of Physical Activity and Micronutrients on Cognitive Performance in Children Aged 6 to 11 Years: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Medicina*, 58(1), 57. <https://doi.org/10.3390/medicina58010057>

Nakahara, N., Matsuyama, Y., Kino, S., Badrakhkhuu, N., Ogawa, T., Moriyama, K., et al. (2020). The consumption of sweets and academic performance among mongolian children. *International Journal of Environmental Research and Public Health*, 17(23), 8912. <https://doi.org/10.3390/ijerph17238912>

Naveed, S., Lakka, T., & Haapala, E. A. (2020). An overview on the associations between health behaviors and brain health in children and adolescents with special reference to diet quality. *International Journal of Environmental Research and Public Health*, 17(3), 953. <https://doi.org/10.3390/ijerph17030953>

Nawabjan, A., & Nazni, P. (2021). Influence of dietary intake and study habits on academic performance in selected residential tribal welfare school students in Salem district of Tamil Nadu. *International Journal of Food and Nutritional Sciences*, 10(2), 30-30. https://doi.org/10.4103/ijfans.ijfans_33_21

Papasideris, M., Safati, A., Ayaz, H., Morita, P., & Hall, P. (2020). Examining the relationship between adolescent health behaviors, brain health, and academic achievement using fNIRS. *medRxiv*. <https://doi.org/10.1101/2020.09.08.20190835>

Pardo, A. & San Martín, R, (2010). *Analisis de datos en Ciencias Sociales y de la Salud II*. Editorial Síntesis.

Qasrawi, R., Halawa, D. A. A., Ayyad, R., Sabah, H. A., Taweeel, H., & Abdeen, Z. (2021). Links between nutrition, life style habits and academic achievement in Palestinian schoolchildren: A cross-sectional study. *Al-Quds University Journal for Academic Research*, 1(1). <https://doi.org/10.47874/2021p6>

Quispe, A., Pinto, D., Huamán, M., Bueno, G., & Valle-Campos, A. (2020). Metodologías cuantitativas: Cálculo del tamaño de muestra con STATA y R. *Revista Del Cuerpo Médico Hospital Nacional Almanzor Aguinaga Asenjo*, 13(1), 78 - 83. <https://doi.org/10.35434/rcmhnaaa.2020.131.627>

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Raschka, S. & Mirjalili, V. (2019). Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow. *Marcombo*

Rosa-Guillamón, A., García-Cantó, E., & Carrillo-López, P.J. (2018). La educación física como programa de desarrollo físico y motor. *Revista Digital de Educación Física*, (52), 105-124.

Rosa-Guillamón, A., López, P. J., Cantó, E., & García, J. E. (2020). La organización y la gestión de la clase de Educación Física. *VIREF*, 9(4), 81-96.

Serra-Majem, L., & Ortiz-Andrellucchi, A. (2018). La dieta mediterránea como ejemplo de una alimentación y nutrición sostenibles: enfoque multidisciplinar. *Nutrición hospitalaria*, 35(SPE4), 96-101. <https://doi.org/10.20960/nh.2133>

Serra-Majem, L., Ribas, L., Ngo, J., Ortega, R. M., García, A., Pérez-Rodríguez, C., & Aranceta, J. (2004). Food, youth and the Mediterranean diet

in Spain. Development of KIDMED, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutrition*, 7(7), 931–935. <https://doi.org/10.1079/PHN2004556>

Si, J., Zhang, H., Zhu, L., & Chen, A. (2021). The relationship between overweight/obesity and executive control in college students: The mediating effect of BDNF and 5-HT. *Life*, 11(4), 313. <https://doi.org/10.3390/life11040313>

Tapia-Serrano, M. A., Esteban-Cornejo, I., Rodriguez-Ayllon, M., Vaquero-Solís, M., Sánchez-Oliva, D., & Sánchez-Miguel, P. A. (2021). Adherence to the Mediterranean diet and academic performance in adolescents: Does BMI status moderate this association? *Clinical Nutrition*. <https://doi.org/10.1016/j.clnu.2020.12.036>

Teixeira, B., Afonso, C., Rodrigues, S., & Oliveira, A. (2021). Healthy and sustainable dietary patterns in children and adolescents: a systematic review. *Advances in Nutrition*. <https://doi.org/10.1093/advances/nmab148>

Tserne, T. A., Borisenkova, M. F., Popov, S. V., Bakutova, L. A., Jongte, L., Trivedi, A. K., et al. (2021). Food Addiction and Weight in Students with High Academic Performance. *Public Health Nutrition*, 1-20. <https://doi.org/10.1017/S1368980021002160>

Vik, F. N., Nilsen, T., & Øverby, N. C. (2022). Aspects of nutritional deficits and cognitive outcomes—Triangulation across time and subject domains among students and teachers in TIMSS. *International Journal of Educational Development*, 89, 102553. <https://doi.org/10.1016/j.ijedudev.2022.102553>

Contact address: Pedro José Carrillo-López. Government of the Canary Islands.. Pº Pérez Valero, 4, 38650 Los Cristianos, Santa Cruz de Tenerife, Spain. E-mail: pcarlop@canariaseducacion.es

Mastery-Approach Goals in Secondary Education Students

Las Metas de Aproximación al Dominio en estudiantes secundaria

<https://doi.org/10.4438/1988-592X-RE-2023-401-589>

Francisco Quiñonez Tapia

<https://orcid.org/0000-0002-7371-0690>

Universidad de Guadalajara (México)

María de Lourdes Vargas Garduño

<https://orcid.org/0000-0002-7384-3098>

Universidad Michoacana de San Nicolás de Hidalgo (México)

Abstract

Mastery-approach goals positively affect students' academic and emotional development. However, these themes have not been studied from the perspective of their relationship with self-efficacy and meaning in life. Accordingly, we aimed at assessing the effects of self-efficacy and meaning in life on mastery-approach goals in secondary education students. For this purpose, we conducted a cross-sectional quantitative study using a sample of 331,431 students. Of those, 170,739 (51.5%) were females, and 160,692 (48.5%) were males. We used data obtained from the Programme for International Student Assessment (PISA) 2018 results for 58 countries. We also used the meaning in life (EUDMO), self-efficacy (RESILIENCE), learning goals (MASTGOAL), and motivation to master tasks (WORKMAST) scales to evaluate the measurement and structural models. Using structural equation modeling, we analyzed the proposed theoretical model and the research data. We found the following: standardized root mean square residual = 0.022, root mean square error of approximation = 0.048 (90% confidence interval: 0.047, 0.048), comparative fit index = 0.971, and Tucker-Lewis index = 0.963. Using multigroup structural equation modeling, we searched for evidence of configural, metric, and scalar invariance of the structural model within the

sample groups by gender and continent of residence. This study found that self-efficacy and meaning in life positively affected students' mastery-approach goals regardless of gender and continent of residence. The findings of this study will equip teachers, educational psychologists, administrators, and policymakers to include work on self-efficacy and meaning in life in secondary school curricula in order to influence secondary education students' development of achievement goals and thus improve their academic performance and well-being and support the United Nations Educational, Scientific and Cultural Organization's related proposals.

Keywords: learning, motivation, goal orientation, adolescence, self-efficacy, learning goals.

Resumen

Las Metas de Aproximación del Dominio tienen efectos positivos en el desarrollo académico y emocional de los estudiantes. Este tipo de temas no han sido estudiados desde su vinculación conjunta con Sentido de Vida y Autoeficacia, por lo que analizamos el efecto que tienen la Autoeficacia y el Sentido de Vida sobre las Metas de Aproximación al Dominio en estudiantes de secundaria. Fue un estudio cuantitativo transversal. Utilizamos una muestra de 331,431 estudiantes, recuperada de los resultados del Programa de Evaluación Internacional de los Alumnos (PISA) 2018, aplicado en 58 países; 170,739 (51.5%) fueron mujeres y 160,692 (48.5%) fueron hombres. Asimismo, utilizamos las escalas de EUDMO (Sentido de Vida) y RESILIENCE (Autoeficacia), MASTGOAL (Metas de Aprendizaje) y WORKMAST (Motivación para Dominar las Tareas) para evaluar los modelos de medida y el modelo estructural. Con el modelado de ecuaciones estructurales analizamos el modelo teórico propuesto y los datos de la investigación; encontramos lo siguientes resultados: SRMR = .022, RMSEA = .048 (90% IC: .047, .048), CFI = .971 y TLI = .963. A partir del Modelado de Ecuaciones Estructurales Multigrupo buscamos las evidencias de invarianza configuracional, métrica y escalar del modelo estructural dentro de los grupos de la muestra por sexo y continente de residencia. La Autoeficacia y el Sentido de Vida tienen un efecto positivo en las Metas de Aproximación al Dominio en estudiantes, y este efecto es invariante por el sexo y continente de residencia de los participantes. Se obtiene conocimiento para que los docentes, psicólogos educativos, directivos y formuladores de políticas públicas educativas incluyan el trabajo de la Autoeficacia y el Sentido de la Vida en los estudiantes de secundaria para incidir en el desarrollo de las metas de logro y así mejorar el rendimiento académico y bienestar de estudiantes de secundaria y apoyar las propuestas de UNESCO.

Palabras clave: aprendizaje, motivación, orientación a la meta, adolescencia, autoeficacia, metas de aprendizaje

Introduction

During adolescence, the stage of life when students attend secondary school, several problems can lead to poor academic achievement, desertion, and school dropout, such as identity crises (Erickson, 2004) and limited development of socioemotional competences including discipline, motivation, and time management, among others (Rodríguez, 2021). Motivation is an impulse that drives adolescents to take action to achieve goals and depends on their biopsychosocial factors, environment, and interests (Castro Castiblanco, Puentes, & Guerrero Cruz, 2019). Although other factors beyond motivation affect school success, motivation considerably influences academic performance. Thus, analyzing achievement goal setting may provide relevant information for designing educational strategies that strengthen mastery-approach goals, especially when associated with self-efficacy and meaning in life given that these elements significantly affect the development of such goals. Therefore, the aim of this study was to assess the effects of self-efficacy and meaning in life on mastery-approach goals in secondary education students. Hence, our study is in line with the United Nations Educational, Scientific and Cultural Organization's (UNESCO) call to "face [the] dual challenge of making good on the unfulfilled promise to ensure the right to quality education for every child, youth and adult and fully realizing the transformational potential of education as a route for sustainable collective futures" (2022, p. 3).

Mastery-Approach Goals

Achievement goal theory proposes to study the goal-setting motivational processes that drive a subject's adaptive or maladaptive behavior in cognitive tasks (Dweck, 1986). In this regard, Elliot (1999) purported that achievement goals allow the subject to generate intrinsic motivations for energizing and guiding cognitive and affective processes based on targeted skills. Achievement goal theory establishes mastery-approach and performance goals according to Dweck (1986); specifically, regarding mastery-approach goals, also known as learning goals, subjects focus on developing competences or understanding something new until they master them, whereas with performance goals, subjects aim at outperforming

others at certain tasks in order to achieve recognition through favorable judgments about their competence and avoid negative judgments; their motive is to demonstrate their ability in front of an audience. Thus, the dichotomous achievement goal model emerged in this context.

The mastery goals Dweck (1986) proposed were adopted in the trichotomous model (Elliot & Harackiewicz, 1996) and bifurcated into performance goals, yielding mastery-approach goals, performance-approach goals, and performance-avoidance goals. Based on this trichotomous model, the 2×2 achievement goal framework (Elliot, 1999) was proposed, comprising mastery-approach goals, mastery-avoidance goals, performance-approach goals, and performance-avoidance goals. In the 2×2 achievement goal framework, Elliot (1999) defined mastery-approach goals as those regarding which subjects strive to develop their abilities until they master the targeted tasks and mastery avoidance goals as those in which subjects strive to avoid losing their abilities (e.g., “striving to avoid leaving a crossword puzzle incomplete”; Elliot & McGregor, 2001, p. 502).

Based on the 2×2 achievement goal framework, the 3×2 achievement goal model was proposed comprising six goals, namely task-approach goals, self-approach goals, other-approach goals, task-avoidance goals, self-avoidance goals, and other-avoidance goals (Elliot et al., 2011). According to Elliot et al. (2011), task-approach goals aim at developing competences in meeting the absolute demands of tasks, whereas self-approach goals aim at developing competences to strengthen personal development, persistence, motivation, and enthusiasm. He also stated that mastery-approach goals encompass task- and self-approach goals and that they are often mixed in daily life. Therefore, the first hypothesis this study proposed was that mastery goals and motivation to master tasks can be explained by a higher, second-order factor termed mastery-approach goals (see Fig. I).

As shown in the literature, mastery-approach goals have a positive effect on students' academic performance (Alhadabi & Karpinski, 2020; Tuominen et al., 2020), as well as on deep (Aydiner-Uygun, 2020), reflective, and integrative (Miller, Fassett, & Palmer, 2021) learning, metacognition (Jaitner et al., 2019), and the ability to transfer knowledge in order to solve new problems (Belenky & Nokes-Malach, 2013).

Mastery-approach goals increase students' positive emotions (Datu, Valdez, & Yang, 2022), decrease their psychological stress, depression, and anxiety levels (Danthon, Mascret, & Cury, 2021), and are correlated with achievement emotions such as interest, enjoyment, hope, and pride (Huang, 2011).

On the one hand, some researchers have reported that females are more likely to develop mastery-approach goals (Camacho, et al., 2022; Nie & Liem, 2013), whereas others have reported that males display higher levels of such goals (Arens & Watermann, Méndez-Giménez, García-Romero, & Cecchini-Estrada, 2018). Other researchers have reported no gendered differences in mastery-approach goal levels (Lochbaum, Zanatta, & Kazak, 2019; Urdan & Kaplan, 2020).

On the other hand, Lochbaum et al. (2019) found that more individualistic countries have higher levels of mastery-approach goals than less individualistic countries. Similarly, Urdan and Kaplan (2020) observed that the culture of the participants had diverse effects on mastery-approach goals, albeit with nonsignificant differences.

Self-Efficacy

Self-efficacy refers to subjects' beliefs regarding their own ability to complete a specific task; these beliefs enable them to organize and execute actions in order to perform the task (Bandura, 1986). Self-efficacy is developed through mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states (Bandura, 1977). Self-efficacy has many positive effects on academic performance (Drago et al., 2018), language proficiency (Wang & Sun, 2020), mathematics learning (Huang, 2016), and creativity (Haase et al., 2018), among others. According to Huang (2013), males have higher self-efficacy than females; however, that scholar reported a rather small difference, and other researchers found no gendered differences in self-efficacy (Assouline et al., 2021). Furthermore, Huang (2013) found no evidence that culture moderated self-efficacy and speculated that perhaps such results were due to low statistical power and to the low diversity of the countries included in the meta-analysis.

Meaning in Life

Meaning in life refers to the degree to which subjects understand their life and give meaning to their self-perception based on a general purpose in life, thereby generating the sense that their life matters (Steger, 2009). Thus, meaning in life has three components: (a) comprehension/coherence, which enables subjects to build coherent meaning frameworks that provide them with expla-

nations for and ascribe meaning to their existence including past, present, and imagined future events (George & Park, 2016); (b) purpose, which enables subjects to identify the desired objects of life and generate directed and motivated behavior to achieve those objects (King & Hicks, 2021); and (c) significance, which generates subjects' belief that their life has a significant impact on reality and will have lasting importance after their death (Martela & Steger, 2016). Positive affects, social connections, self-connections, religion and worldview, the ability to visualize the past and the future, and awareness of mortality all comprise meaning in life (King & Hicks, 2021).

Meaning in life has a positive effect on adaptability, professional self-efficacy (Yuen & Chan, 2022), life satisfaction (Heng et al., 2020), and academic and personal self-efficacy (Yuen & Datu, 2021). In students, meaning in life has been positively correlated with motivation, positive affects, subjective well-being, and high grades (Bailey & Phillips, 2016).

Geng et al. (2022) and Yuen and Chan (2022) found no significant differences in meaning in life levels between males and females. However, Hamama and Hamama-Raz (2021) reported that females have higher levels of meaning in life than males. Lastly, regarding cultural differences, Heng et al. (2020) reported that Israeli students have higher levels of meaning in life than Singaporean students.

Research Model

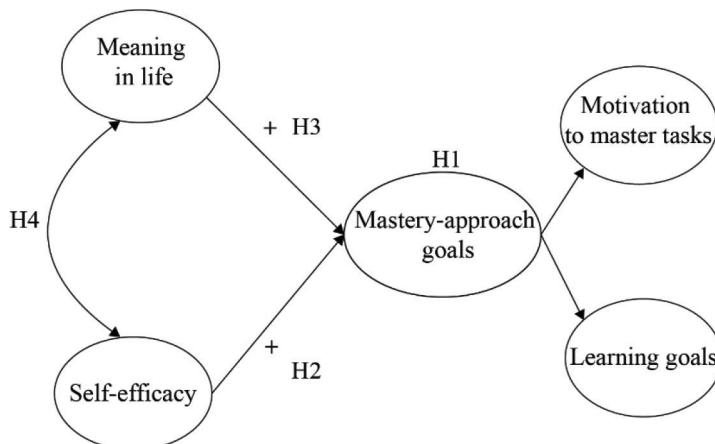
Predictors of mastery-approach goals have been reported in the literature, albeit overlooking their combined interaction with self-efficacy and meaning in life. Consequently, we propose a structural model for studying the relationship of mastery-approach goals with self-efficacy and meaning in life, involving two causal relationships and one correlational relationship (see Fig. I).

The first causal relationship refers to the effect of self-efficacy on mastery-approach goals and is based on previous findings showing that self-efficacy has positive effects on (Alhadabi & Karpinski, 2020; Ariani, 2022) and is correlated with (Huang, 2016) students' mastery-approach goals. Additionally, gender is a moderator of the relationship between self-efficacy and mastery-approach goals (Huang, 2016). Based on the above, our second hypothesis was that self-efficacy would have a positive effect on mastery-approach goals among secondary education students.

As the second causal relationship, we proposed that meaning in life would affect mastery-approach goals (see Fig. I). We did not find studies on the relationship between meaning in life and mastery-approach goals. However, this relationship is worthy of analysis because meaning in life is a component of subjective well-being (Steger, 2017), and increasing meaning in life increases subjective well-being (Li et al., 2021). Accordingly, a positive correlation of subjective well-being with mastery-approach goals has been found in students (Ariani, 2022; Li & Zhao et al., 2021). Based on the above, our third hypothesis was that meaning in life would have a positive effect on mastery-approach goals in secondary education students.

In the model developed in this study, we proposed a correlation of meaning in life with self-efficacy because the latter is a positive predictor of the former (Cheng, Chen, & Zhang, 2021), and vice-versa (Rush et al., 2021). However, we did not find literature on the effects of gender and place of residence on the relationship between the aforementioned variables (see Fig. I). Consequently, our fourth hypothesis was that self-efficacy would show a positive correlation with meaning in life among secondary education students.

FIGURE I. Visualization of the proposed model



Source: The authors.

Lastly, we analyzed the proposed model's invariance within groups by gender (males and females) and continent of residence (America, Asia, and Europe). As shown above, the results regarding gendered and cultural differences with respect to mastery-approach goals, self-efficacy, and meaning in life were heterogeneous. Thus, our fifth hypothesis was that the effect of self-efficacy and meaning in life on mastery-approach goals would not vary according to subjects' gender or continent of residence.

Method

Sample and Procedure

The Organisation for Economic Co-operation and Development's (OECD, 2018a) Programme for International Student Assessment (PISA) evaluates students from 79 countries at the end of compulsory education. We retrieved the data from PISA 2018 and selected a student body with complete scores for the scales used to measure meaning in life (EUDMO), self-efficacy (RESILIENCE), learning goals (MASTGOAL), and motivation to master tasks (WORKMAST; OECD, 2018b). This study's final sample comprised 331,431 students, of whom 170,739 (51.5%) were females and 160,692 (48.5%) were males. Participant age ranged from 15 years to 16 years (mean [M] = 15.79, standard deviation [SD] = 0.29). Table I outlines the distribution of participants by continent and country of residence.

TABLE I. Participants' continent and country of residence

Continent/Country	<i>n</i>	%	Continent/Country	<i>n</i>	%
America	46,838	14.1	Europe	166,333	50.2
Argentina	8,302	2.5	Albania	5,171	1.6
Brazil	6,618	2.0	Baku (Azerbaijan)	3,279	1.0
Chile	5,570	1.7	Austria	5,498	1.7
Colombia	5,564	1.7	Bosnia and Herzegovina	4,729	1.4
Costa Rica	5,715	1.7	Bulgaria	3,358	1.0
Dominican Republic	1,870	0.6	Belarus	5,018	1.5

(Continued)

TABLE I. Participants' continent and country of residence (Continued)

Continent/Country	<i>n</i>	%	Continent/Country	<i>n</i>	%
Mexico	4,621	1.4	Croatia	5,498	1.7
Panama	2,243	0.7	Estonia	4,664	1.4
Peru	3,322	1.0	France	4,773	1.4
Uruguay	3,013	0.9	Germany	3,200	1.0
			Greece	5,250	1.6
Asia	118,260	35.7	Hungary	4,395	1.3
Brunei Darussalam	5,334	1.6	Iceland	2,604	0.8
Chinese Taipei	6,803	2.1	Ireland	4,797	1.4
Hong Kong	5,461	1.6	Italy	8,960	2.7
Indonesia	10,708	3.2	Kosovo	3,616	1.1
Kazakhstan	15,386	4.6	Latvia	4,490	1.4
Jordan	7,054	2.1	Lithuania	5,642	1.7
Korea	6,336	1.9	Malta	2,715	0.8
Macao	3,643	1.1	Moldavia	4,635	1.4
Malaysia	5,709	1.7	Montenegro	4,841	1.5
Philippines	5,851	1.8	Poland	5,096	1.5
Russia	6,019	1.8	Portugal	4,978	1.5
Saudi Arabia	4,651	1.4	Romania	4,366	1.3
Thailand	7,859	2.4	Serbia	4,477	1.4
United Arab Emirates	15,112	4.6	Slovakia	4,676	1.4
Turkey	6,055	1.8	Slovenia	5,296	1.6
Moscow Oblast (RUS)	1,630	0.5	Spain	28,358	8.6
Tatarstan (RUS)	4,649	1.4	Switzerland	4,211	1.3
			Ukraine	5,347	1.6
			United Kingdom	2,395	0.7

Source. Compiled by authors

Instruments

EUDMO, RESILIENCE, WORKMAST and MASTGOAL (OECD, 2018c, 2019b) are scored on a Likert-type scale. EUDMO, RESILIENCE, and WORKMAST are 4-point scales, consisting of items with four response options (ranging from 1 = *strongly disagree* to 4 = *strongly agree*), and MASTGOAL is a 5-point scale (ranging from 1 = *Not at all true of me* to 5 = *Extremely true of me*). For the scales measuring meaning in life (EUDMO) and self-efficacy (RESILIENCE), scores range from 3 points to 12 points; for the scale measuring learning goals (MASTGOAL), scores are between 5 points and 20 points, and for motivation to master tasks (WORKMAST), scores range from 4 points to 12 points. In the obtained data, the reported individual total scores were probability-weighted estimates transformed to have a mean of 0.0 and a standard deviation of 1.0 across OECD countries (2019b, s.f.).

EUDMO (OECD, 2018c), which measures meaning in life, consists of three items on students' sense of life, meaning of life, and purpose of life, respectively (e.g., "My life has a clear meaning or purpose."), and positive values indicate greater meaning in life than the average student in OECD countries (2019b). Reliability analysis of the study participants' scores on the EUDMO scale yielded a Cronbach's alpha of 0.85. To assess self-efficacy, we used RESILIENCE (OECD, 2019b), which consists of five items (e.g., "I usually manage one way or another."). Positive values indicate greater self-efficacy than the average student in OECD countries (2019b). Reliability analysis of the study participants' scores on the RESILIENCE scale yielded a Cronbach's alpha of 0.78.

MASTGOAL (OECD, 2018c) measures learning goals and consists of three items concerning students' mastery-approach orientation toward learning goals (e.g., "My goal is to completely master the material presented in my classes."). Positive values indicate more ambitious learning goals than the average student across OECD countries (2019b). Reliability analysis of the study participants' scores on the MASTGOAL scale yielded a Cronbach's alpha of 0.87. Lastly, WORKMAST (OECD, 2019b) was used to assess motivation to master tasks. This scale consists of three items covering students' motivation to work as well as their achievement (e.g., "Once I start a task, I persist until it is finished."). Positive values indicate greater motivation to master tasks than the average student across OECD countries (2019b). Reliability analysis of the study participants' scores on the WORKMAST scale yielded a Cronbach's alpha of 0.78.

Data Analysis

We calculated the mean, standard deviation, and percentage of the data to determine the descriptive statistics. We also performed Student's t-test to compare participants' scores on each scale by gender, and we calculated Cohen's *d* (Cohen, 1988) to assess effect size differences between the study groups. Additionally, we performed a one-way analysis of variance (ANOVA) (Christensen, 2016) to compare samples grouped by continent of residence and a two-way ANOVA to examine the interaction between gender and continent of residence with respect to participants' scores in order to determine those factors' relationship to the study variables. We also analyzed Cohen's *f* (Cohen, 1988) to determine the effect size between groups compared with the ANOVA. We calculated the statistical power (1- β) to confirm that the null hypothesis was correctly rejected, based on between-group score comparison (Cohen, 1988). Following Cohen (1992), we assessed and interpreted the effect size of the proposed inter-group comparisons with Cohen's *d* values of 0.20 (small), 0.50 (medium), and 0.80 (large) and with Cohen's *f* values of 0.10 (small), 0.25 (medium), and 0.40 (large). We also set a cut-off point of $1-\beta \geq 0.80$ (Cohen, 1992) for statistical power.

We analyzed the study hypotheses via structural equation modeling (Hancock & Mueller, 2013). We evaluated and interpreted the goodness-of-fit between the data and the proposed model using the following statistical indices and cut-off points: chi-square (χ^2), its degrees of freedom (*df*), and its level of significance; standardized root mean squared residual (SRMR = 0.08); root mean square error of approximation (RMSEA ≤ 0.06); comparative fit index (CFI ≥ 0.95), and Tucker-Lewis indices (TLI ≥ 0.95 ; Hu & Bentler, 1999).

To gather evidence of invariance, we performed multigroup analysis of the proposed model and analyzed the configural, metric, and scalar invariance (Vandenberg & Lance, 2000). To interpret the model invariance evidence for different sample groups, we analyzed changes in the chi square and CFI increment. We established that a significant chi-square value ($p = 0.05$) or a CFI increment equal to or less than -0.01 indicated model invariance (Cheung & Rensvold, 2002).

Through confirmatory factor analysis (CFA), we estimated measurement models to assess the evidence of construct validity. Using analysis of variance extracted (AVE), we examined the evidence of convergent validity in the models and calculated the composite reliability (CR). In line

with Fornell and Larcker (1981), we assessed the evidence of convergent validity in the models by setting the following values: CR ≥ 0.60 and AVE ≥ 0.50 . We used SPSS 27, AMOS 27, and GPower 3.1 to perform the tests.

Results

Descriptive Statistics and Between-Group Comparisons

The results of analysis of the mean and standard deviation and inter-group comparison of the study participants' scores by scale and gender are outlined in Table II, showing that females scored significantly ($p < 0.001$) higher than males on learning goals ($M = 0.217$, $SD = 1.022$) and motivation to master tasks ($M = 0.262$, $SD = 0.979$), but the effect size was small ($d = 0.18$ – 0.19). Conversely, males scored significantly ($p < 0.001$) higher than females on meaning in life ($M = 0.218$, $SD = 0.988$) and self-efficacy ($M = 0.110$, $SD = 1.039$), but the effect size was nonsignificant ($d = 0.01$ – 0.07).

As shown in Table III, American participants scored higher on meaning in life ($M = 0.291$, $SD = 0.994$), self-efficacy ($M = 0.239$, $SD = 1.041$), learning goals ($M = 0.347$, $SD = 1.057$), and motivation to master tasks ($M = 0.361$, $SD = 1.029$). Table 3 also highlights significant differences ($p < 0.001$) in the study variables according to the participants' continent of residence but nonsignificant effect sizes ($f < 0.10$) for scores in

TABLE II. Means and standard deviations and gendered comparison of study participants' scores on each scale

Variables	Females		Males		t	gI	95% CI	d	β
	M	SD	M	SD					
Meaning in Life	0.142	0.959	0.218	0.988	-22.62***	331,429	[-0.08, -0.07]	0.07	1
Self-efficacy	0.091	0.973	0.110	1.039	-5.46***	326,293	[-0.03, -0.01]	0.01	1
Learning Goals	0.217	1.022	0.030	1.061	51.54***	328,232	[0.18, 0.19]	0.18	1
Motivation to Master Tasks	0.262	0.979	0.076	1.022	53.55***	331,439	[0.18, 0.19]	0.19	1

Note. CI: confidence interval; *** $p < 0.001$

Source: Compiled by authors

TABLE III. Means, standard deviations, and comparison of the participants' response rates for each scale by continent of residence

Variables	America		Asia		Europe		F (2)	β	
	M	SD	M	SD	M	SD			
Meaning in Life	0.291	0.994	0.240	0.946	0.103	0.981	1048***	0.08	1
Self-efficacy	0.239	1.041	0.024	0.999	0.115	0.995	808***	0.07	1
Learning Goals	0.347	1.057	0.232	0.104	-0.009	1.022	3125***	0.14	1
Motivation to Master Tasks	0.361	1.029	0.188	1.011	0.107	0.986	1200***	0.08	1

Note. *** $p < 0.001$

Source: Compiled by authors

meaning in life, self-efficacy, and motivation to master tasks, as well as a small effect size ($f = 0.14$) for scores on learning goals. In all groups, the statistical power was 1.

The results of the two-way ANOVA showed significant differences in the participants' scores by gender and continent of residence, with a nonsignificant effect size with respect to the scales as follows: Meaning in Life: $F(2) = 193$, $p < 0.001$, $f = 0.03$, $\beta = 1$; Self-efficacy: $F(2) = 78$, $p < 0.001$, $f < 0.001$, $\beta = 1$; Learning Goals: $F(2) = 17$, $p < 0.001$, $f < 0.001$, $\beta = 1$; and Motivation to Master Tasks: $F(2) = 19$, $p < 0.001$, $f < 0.001$, $\beta = 1$.

Measurement Model

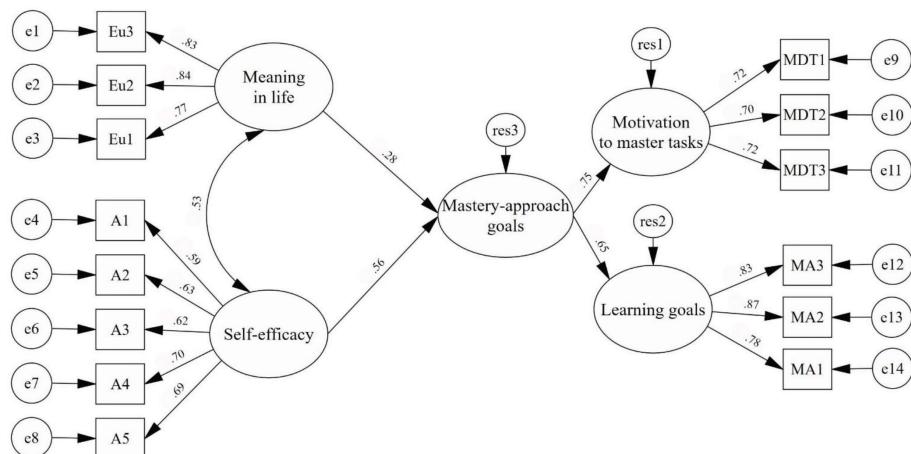
The measurement models were estimated using CFA, CR, and AVE; the results showed evidence of construct and convergent validity regarding the study variables. Analysis of the self-efficacy measurement model yielded the following values: SRMR = 0.031, RMSEA = 0.095 (90% CI: 0.094, 0.096), CFI = 0.927, TLI = 0.963, CR = 0.782, and AVE = 0.419, with the indicators' standardized factor loadings ranging from 0.60 to 0.70. Analysis of the meaning in life measurement model yielded the following values: SRMR = 0.011, RMSEA = 0.084 (90% CI: 0.056, 0.058), CFI = 0.995, TLI = 0.984, CR = 0.840, and AVE = 0.637, with the indicators' standardized factor loadings ranging from 0.78 to 0.83.

Additionally, analysis of the second-order measurement model for mastery-approach goals yielded the following values: SRMR = 0.025, RMSEA = 0.057 (90% CI: 0.081, 0.087), CFI = 0.989, and TLI = 0.979. For motivation to master tasks, the analysis showed that CR = 0.757 and AVE = 0.509, with the indicators' standardized factor loadings ranging from 0.69 to 0.73. For learning goals, the values were CR = 0.867 and AVE = 0.685, with the indicators' standardized factor loadings ranging from 0.77 to 0.88. Consequently, we accepted Hypothesis 1.

Structural Model

The model explaining mastery-approach goals based on self-efficacy and meaning in life among secondary education students (see Fig. II) fit the data very well: SRMR = 0.022, RMSEA = 0.048 (90% CI: 0.047, 0.048), CFI = 0.971, and TLI = 0.963. Fig. II shows that the latent variables meaning in life ($\beta = 0.28, p < 0.001$) and self-efficacy ($\beta = 0.56, p < 0.001$)

FIGURE II. Structural equation model explaining mastery-approach goals in secondary education students



Note. This model explains mastery-approach goals in secondary education students based on self-efficacy and meaning in life. The statistics are standardized regression coefficients. All path coefficients and correlations in the model are significant ($p < .001$).
Source: Compiled by authors

had significant and positive effects on mastery-approach goals, thus confirming Hypotheses 2 and 3. Furthermore, Fig. II shows significant and positive correlations between meaning in life and self-efficacy ($\beta = 0.53$, $p < 0.001$). Therefore, we accepted Hypothesis 4. Lastly, the explained variance of mastery-approach goals based on the latent variables self-efficacy and meaning in life was 56.3% ($R^2 = 0.563$).

Invariance

Table IV presents the results of analysis of invariance in the proposed model for explaining mastery-approach goals based on self-efficacy and meaning in life in secondary education students by gender and continent of residence. The chi-squared difference test results showed the structural model's lack of invariance in the compared groups ($p < 0.05$). However, the chi-square difference test is sensitive to sample size. Accordingly, Cheung and Rensvold (2002) have suggested evaluating model invariance according to the CFI increment. Such analysis yielded evidence of configural ($CFI > 0.95$), metric, and scalar invariance ($\Delta CFI \leq -0.01$) with respect to the model under analysis for different groups of secondary education students (see Table IV). Therefore, we accepted Hypothesis 5.

TABLE IV. Structural equation model's goodness-of-fit indices by group

Models	χ^2	gl	$\Delta \chi^2$	CFI	ΔCFI	RSMEA [90% CI]
Sex						
M0. Configural invariance	52,119	144		0.972		0.033[0.033, 0.033]
M1. Metric invariance	54,853	154	2,737***	0.970	-0.002	0.033[0.033, 0.033]
M2. Scalar invariance	55,100	157	247***	0.970	0.000	0.032[0.032, 0.033]
Continent						
M0. Configural invariance	55,216	216		0.970		0.028[0.028, 0.028]
M1. Metric invariance	57,163	236	1,947***	0.969	-0.001	0.027[0.027, 0.027]
M2. Scalar invariance	58,185	242	1,022***	0.969	0.000	0.027[0.027, 0.027]

Note. To determine the invariance measures, the models were compared as follows: M1-M0 and M2-M1; *** $p < 0.001$
Source: Compiled by authors

Discussion and Conclusion

In this study, we aimed at explaining mastery-approach goals based on self-efficacy and meaning in life among secondary education students from 58 countries evaluated in the 2018 PISA. For this purpose, we performed an analysis by gender and culture to assess whether these factors affected the relationships between those variables. All hypotheses were accepted. Self-efficacy and meaning in life were found to be positively correlated with and showed a positive effect on mastery-approach goals in secondary education students, and these were consistent between groups of students divided by gender and continent of residence. We also found significant differences in the levels of mastery-approach goals, self-efficacy, and meaning in life between males and females and between participants living in America, Asia, and Europe, albeit with negligible to small effect sizes.

Self-efficacy had a positive effect on mastery-approach goals in secondary education students (Jiang et al., 2014), and the effect did not vary with gender (Huang, 2016). Moreover, self-efficacy was a consistent predictor of mastery-approach goals. We found insignificant differences in self-efficacy between males and females and between students living in America, Asia, and Europe (Assouline et al., 2021; Huang, 2013). Gendered differences in the levels of mastery-approach goals were small (Nie & Liem, 2013), as were differences by continent of residence (Urdan & Kaplan, 2020). Considering the above, self-efficacy may be a component of mastery-approach goals.

Meaning in life had a positive effect on mastery-approach goals in secondary education students. Although we did not find empirical studies on the relationship between meaning in life and mastery-approach goals, we are able to explain our results theoretically because meaning in life and mastery-approach goals share a fundamental component: the ability to set objectives in life. Therefore, we consider meaning in life to be theoretically related to mastery-approach goals.

On the one hand, meaning in life enables individuals to establish a general purpose in life, whereas mastery-approach goals enable them to establish a specific purpose that directs their behavior in a particular direction. On the other hand, meaning in life is considered to be a dimension of subjective well-being (Steger, 2017), and our results have demonstrated that subjective well-being can be used as a predictor of

mastery-approach goals (Kaplan & Maehr, 1999; Zhou, Huebner, & Tian, 2020). However, we did not find significant gendered differences in meaning in life levels (Hamama & Hamama-Raz, 2021). Therefore, our results corroborate the findings of Geng et al. (2022), Yuen and Chan (2022), and Hamama and Hamama-Raz (2021), who also did not find significant gendered differences.

Our results support empirical evidence of a correlation between self-efficacy and meaning in life (Cheng et al., 2020; Cheng et al., 2021; Rush et al., 2021). This result confirms the importance of self-efficacy, not only for setting mastery-approach goals but also for strengthening meaning in life among secondary education students. Therefore, fostering meaning in life among students will strengthen self-efficacy, and vice versa, which will enable them to develop mastery-approach goals and improve learning.

This study's main strength is that it used the largest sample that has been reported thus far in research related to achievement goal theory. The sample included students from 58 countries and three continents, which makes it possible to generalize the results across all the participating nations. Additionally, different invariance analyses of the proposed structural model were performed according to participants' gender and continent of residence.

Nevertheless, this study had some limitations. The research addresses achievement goal theory as a whole but focused on mastery-approach goals, disregarding mastery-avoidance goals, performance-approach goals, and performance-avoidance goals. By including these three latter types of goals in future research, we expect to understand the practical implications associated with students' academic performance in order to propose teaching and learning strategies that will highlight the emotional implications of achievement goal theory.

Another limitation was the use of a self-report instrument, which can generate social desirability bias in the participants' responses and hinder the replicability of the study in specific populations. To overcome this limitation, we propose taking up the topic of mastery-approach goals in research with an experimental design that will allow us to control for variables and bias; further, a qualitative perspective will enable us to analyze in depth the psychological factors that affect subjects' establishment of the abovementioned goals.

Another limitation was that this study only included students aged between 15 years and 16 years; this was so because of the nature of the

data source. The participants' small age range prevented the inclusion of all secondary school students and means that the study generated knowledge pertaining to a very specific age group instead of contributing more to research on the development of mastery-approach goals throughout the subjects' lifetime. To overcome this limitation, we must understand the dynamics of master-approach goals at different stages of human development in order to inform intervention strategies tailored to suit various subjects' current stage of development. In light of the limitation, the results of this study must be judiciously applied to populations of other ages since age is a modifying factor of self-efficacy levels (Huang, 2013; Mozahem, Boulad, & Ghanem, 2021) and mastery-approach goals (Méndez-Giménez et al., 2018). Lastly, because we did not find previous studies on the effect of meaning in life on mastery-approach goals, we suggest further exploring the relationship between these variables and expanding the inquiry with reference to subjective well-being theories; such an undertaking would replicate our results in specific populations and contribute to the development of achievement goal theory.

The results of the present study have contributed to the development of achievement goal theory by providing evidence of predictors of mastery-approach goals in the combined dimensions of self-efficacy and meaning in life among secondary education students. Regarding practical implications, this study's results can serve as a reference for secondary school teachers to design and implement activities and teaching and learning environments aimed at strengthening students' self-efficacy and meaning in life. Thus, students will be guided toward establishing and achieving mastery-approach goals to improve their academic performance and school well-being. Moreover, educational psychologists will be able to draw upon scientific evidence to support the inclusion of themes related to self-efficacy and meaning in life in individual and group intervention programs aimed at improving students' academic performance and subjective well-being. Additionally, this study equips school directors and education policymakers with relevant scientific knowledge to promote educational quality by designing and implementing policies and programs geared toward students' psychological development based on self-efficacy and meaning in life.

In conclusion, we have provided evidence of the importance of self-efficacy and meaning in life in secondary education students' development of mastery-approach goals. All our hypotheses were supported

since students' self-efficacy and meaning in life showed a positive effect on mastery-approach goals, and the effect did not vary with gender or continent of residence. Thus, teachers, educational psychologists, school directors, and education policymakers are advised to include self-efficacy and meaning in life in secondary education curricula in order to promote the development of mastery-approach goals and improve students' academic performance and well-being, thereby helping to tackle the educational challenges UNESCO (2022) has flagged.

Bibliographic References

Alhadabi, A., & Karpinski, A. (2020). Grit, self-efficacy, achievement orientation goals, and academic performance in University students. *International Journal of Adolescence and Youth*, 25(1), 519–535. <https://doi.org/10.1080/02673843.2019.1679202>

Ariani, D. (2022). How achievement goals affect students' well-being and the relationship model between achievement goals, academic self-efficacy and affect at school. *Journal of Applied Research in Higher Education*, 14(1), 111–134. <https://doi.org/10.1108/JARHE-08-2020-0273>

Arens, A. K., & Watermann, R. (2021). Students' achievement goals and beliefs of causes of success: Temporal relations and gender differences. *Contemporary Educational Psychology*, 64, 101941. <https://doi.org/10.1016/j.cedpsych.2020.101941>

Assouline, S. G., Mahatmya, D., Ihrig, L., & Lane, E. (2021). High-achieving rural middle-school students' academic self-efficacy and attributions in relationship to gender. *High Ability Studies*, 32(2), 143–169. <https://doi.org/10.1080/13598139.2020.1740582>

Aydiner-Uygun, M. (2020). Achievement goal orientations of students studying instrument education as predictors of their learning approaches. *Music Education Research*, 22(2), 130–144. <https://doi.org/10.1080/14613808.2020.1713735>

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, 84(4), 139–161. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall series in social learning theory. Prentice-Hall, Inc.

Bailey, T. H., & Phillips, L. J. (2016). The influence of motivation and adaptation on students' subjective well-being, meaning in life and academic performance. *Higher Education Research & Development*, 35(2), 201-216. <https://doi.org/10.1080/07294360.2015.1087474>

Belenky, D., & Nokes-Malach, T. (2013). Mastery-approach goals and knowledge transfer: An investigation into the effects of task structure and framing instructions. *Learning and Individual Differences*, 25, 21-34. <https://doi.org/10.1016/j.lindif.2013.02.004>

Camacho, A., Alves, R. A., Daniel, J. R., De Smedt, F., & Van Keer, H. (2022). Structural relations among implicit theories, achievement goals, and performance in writing. *Learning and Individual Differences*, 100, 102223. <https://doi.org/10.1016/j.lindif.2022.102223>

Castro Castiblanco Y., Puentes, A. & Guerrero Cruz, E. (2019). *Factores que intervienen en la motivación durante la adolescencia y su influencia en el ámbito escolar*. Politécnico Grancolombiano. <https://tinyurl.com/mrj76nxs>

Cheng, L., Chen, Q., & Zhang, F. (2021). Mediating effects of meaning in life on the relationship between general self-efficacy and nursing professional commitment in nursing students A structural equation modeling approach. *Medicine*, 100(29), 1-7. <https://doi.org/10.1097/MD.00000000000026754>

Cheng, L., Ye, Y., Zhong, Z., Zhang, F., Hu, X., Cui, R., & Chen, Q. (2020). Mediating effects of general self-efficacy on the relationship between the source of meaning in life and prosocial behaviours in vocational college nursing students: A cross-sectional study. *PloS One*, 15(12), 1-13. <https://doi.org/10.1371/journal.pone.0243796>

Cheung, G., & Rensvold, R. (2002). Evaluating goodness-of-fit Indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Christensen, R. (2016). *Analysis of Variance, Design, and Regression: Linear Modeling for Unbalanced Data* (2a. ed.). CRC; Taylor & Francis Group; Chapman & Hall Book

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2a. ed.). Routledge.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037/0033-2909.112.1.155>

Danthony, S., Mascret, N., & Cury, F. (2021). The relationships between the 3 x 2 achievement goal model and test anxiety in Physical Education. *European Physical Education Review*, 27(3), 559–573. <https://doi.org/10.1177/1356336X20971325>

Datu, J., Valdez, J., & Yang, W. (2022). The academically engaged life of mastery-oriented students: Causal ordering among positive emotions, mastery-approach goals, and academic engagement. *Revista de Psicodidáctica*, 27(1), 1–8. <https://doi.org/10.1016/j.psicod.2021.02.001>

Drago, A., Rheinheimer, D. C., & Detweiler, T. N. (2018). Effects of locus of control, academic self-efficacy, and tutoring on academic performance. *Journal of College Student Retention: Research, Theory & Practice*, 19(4), 433-451. <https://doi.org/10.1177/1521025116645602>

Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040–1048. <https://doi.org/10.1037/0003-066X.41.10.1040>

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169–189. https://doi.org/10.1207/s15326985ep3403_3

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of personality and social psychology*, 70(3), 461-475. <https://doi.org/10.1037/0022-3514.70.3.461>

Elliot, A. J., & McGregor, H. A. (2001). A 2 X 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., 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>

Erikson, E. (2004). *Sociedad y adolescencia*. Siglo XXI.

Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement Error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>

George, L. S., & Park, C. L. (2016). Meaning in life as comprehension, purpose, and mattering: Toward Integration and New Research Questions. *Review of General Psychology*, 20(3), 205-220. <https://doi.org/10.1037/gpr0000077>

Geng, J., Wang, Y., Wang, P., Zeng, P., & Lei, L. (2022). Gender Differences between Cyberbullying Victimization and Meaning in Life: Roles of

Fatalism and Self-Concept Clarity. *Journal of Interpersonal Violence*, 37(19-20), NP17157-NP17181. <https://doi.org/10.1177/08862605211028285>

Haase, J., Hoff, E. V., Hanel, P. H. P., & Innes-Ker, A. (2018). A Meta-Analysis of the Relation between Creative Self-Efficacy and Different Creativity Measurements. *Creativity Research Journal*, 30(1), 1-16. <https://doi.org/10.1080/10400419.2018.1411436>

Hamama, L., & Hamama-Raz, Y. (2021). Meaning in Life, Self-Control, Positive and Negative Affect: Exploring Gender Differences Among Adolescents. *Youth & Society*, 53(5), 699-722. <https://doi.org/10.1177/0044118X19883736>

Hancock, G., & Mueller, R. (2013). *Structural equation modeling: A second course*. (2nd ed.). Information Age Publishing.

Heng, M. A., Fulmer, G. W., Blau, I., & Pereira, A. (2020). Youth purpose, meaning in life, social support and life satisfaction among adolescents in Singapore and Israel. *Journal of Educational Change*, 21(2), 299-322. <https://doi.org/10.1007/s10833-020-09381-4>

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>

Huang, C. (2011). Achievement goals and achievement emotions: a meta-analysis. *Educational Psychology Review*, 23(3), 359-388. <https://doi.org/10.1007/s10648-011-9155-x>

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, 28(1), 1-35. <https://doi.org/10.1007/s10212-011-0097-y>

Huang, C. (2016). Achievement goals and self-efficacy: A meta-analysis. *Educational Research Review*, 19, 119-137. <https://doi.org/10.1016/j.edurev.2016.07.002>

Jaitner, D., Rinas, R., Becker, C., Niermann, C., Breithecker, J., & Mess, F. (2019). Supporting subject justification by educational psychology: A Systematic Review of Achievement Goal Motivation in School Physical Education. *Frontiers in Education*, 4, 1-25. <https://doi.org/10.3389/feduc.2019.00070>

Jiang, Y., Song, J., Lee, M., & Bong, M. (2014). Self-efficacy and achievement goals as motivational links between perceived contexts and achievement. *Educational Psychology*, 34(1), 92-117. <https://doi.org/10.1080/01443410.2013.863831>

Kaplan, A., & Maehr, M. (1999). Achievement goals and student well-being. *Contemporary Educational Psychology*, 24(4), 330–358. <https://doi.org/10.1006/ceps.1999.0993>

King, L. A., & Hicks, J. A. (2021). The science of meaning in life. *Annual Review of Psychology*, 72(1), 561-584. <https://doi.org/10.1146/annurev-psych-072420-122921>

Li, J., Dou, K., & Liang, Y. (2021). The relationship between presence of meaning, search for meaning, and subjective well-being: A Three-Level Meta-Analysis Based on the Meaning in Life Questionnaire. *Journal of Happiness Studies*, 22(1), 467–489. <https://doi.org/10.1007/s10902-020-00230-y>

Li, Q., Zhao, J., Tian, J., Sun, T., Zhao, C., Guo, H., Zhu, L., Gao, R., Yang, L., Cao, D., & Zhang, S. (2021). The association among achievement goal orientations, academic performance, and academic well-being among Chinese medical students: a cross-sectional study. *Frontiers in Psychology*, 12, 1-10. <https://doi.org/10.3389/fpsyg.2021.694019>

Lochbaum, M., Zanatta, T., & Kazak, Z. (2019). The 2 × 2 achievement goals in sport and physical activity contexts: A meta-analytic test of context, gender, culture, and socioeconomic status differences and analysis of motivations, regulations, affect, effort, and physical activity correlates. *European Journal of Investigation in Health, Psychology and Education*, 10(1), 173-205. <https://doi.org/10.3390/ejihpe10010015>

Martela, F., & Steger, M. F. (2016). The three meanings of meaning in life: Distinguishing coherence, purpose, and significance. *The Journal of Positive Psychology*, 11(5), 531-545. <https://doi.org/10.1080/17439760.2015.1137623>

Méndez-Giménez, A., García-Romero, C., & Cecchini-Estrada, J. A. (2018). Metas de logro 3x2, amistad y afecto en educación física: diferencias edad-sexo. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 18(72), 637. <https://doi.org/10.15366/rimcafd2018.72.003>

Miller, A., Fassett, K., & Palmer, D. (2021). Achievement goal orientation: A predictor of student engagement in higher education. *Motivation and Emotion*, 45(3), 327–344. <https://doi.org/10.1007/s11031-021-09881-7>

Mozahem, N. A., Boulad, F. M., & Ghanem, C. M. (2021). Secondary school students and self-efficacy in mathematics: Gender and age differences. *International Journal of School & Educational Psychology*, 9(sup1), S142-S152. <https://doi.org/10.1080/21683603.2020.1763877>

Nie, Y., & Liem, G. A. D. (2013). Extending antecedents of achievement goals: The double-edged sword effect of social-oriented achievement motive and gender differences. *Learning and Individual Differences*, 23, 249-255. <https://doi.org/10.1016/j.lindif.2012.10.006>

Organisation for Economic Co-operation and Development (s.f). *PISA 2018 Technical Report*. <https://www.oecd.org/pisa/data/pisa2018-technical-report/>

Organisation for Economic Co-operation and Development (2018a). *PISA for Development Assessment and Analytical Framework. Advance online publication*. <https://doi.org/10.1787/9789264305274-en>

Organisation for Economic Co-operation and Development (2018b). *PISA 2018 Database [Data file]*. https://webfs.oecd.org/pisa2018/SPSS_STU_QQQ.zip

Organisation for Economic Co-operation and Development (2018c). *Student Questionnaire for PISA 2018*. <https://www.oecd.org/pisa/publications/Mexico.zip>

Organisation for Economic Co-operation and Development (2019a). *PISA 2018 Results (Volume I). Advance online publication*. <https://doi.org/10.1787/5f07c754-en>

Organisation for Economic Co-operation and Development (2019b). *PISA 2018 Results (Volume III). Advance online publication*. <https://doi.org/10.1787/acd78851-en>

Rodríguez L. (2021). *Los retos y oportunidades de la educación secundaria en América Latina y el Caribe durante y después de la pandemia*. Naciones Unidas. <https://tinyurl.com/4yvxa59p>

Rush, C., Hooker, S., Ross, K., Frers, A., Peters, J., & Masters, K. (2021). Brief report: Meaning in life is mediated by self-efficacy in the prediction of physical activity. *Journal of Health Psychology*, 26(5), 753-757. <https://doi.org/10.1177/1359105319828172>

Steger, M. (2009). Meaning in Life. En S. Lopez and C. Snyder (Ed.), *The Oxford Handbook of Positive Psychology* (2a. ed.), (pp. 678-688). Oxford University Press

Steger, M. (2017). Meaning in Life and Wellbeing. En M. Slade, L. Oades & A. Jarden (Eds.), *Wellbeing, Recovery and Mental Health* (pp. 75-85). Cambridge University Press.

UNESCO (2022). *Reimaginar juntos nuestros futuros. Un Nuevo contrato social para la educación*. https://unesdoc.unesco.org/ark:/48223/pf0000379381_spa

Urdan, T., & Kaplan, A. (2020). The origins, evolution, and future directions of achievement goal theory. *Contemporary Educational Psychology*, 61, 101862. <https://doi.org/10.1016/j.cedpsych.2020.101862>

Tuominen, H., Niemivirta, M., Lonka, K., & Salmela-Aro, K. (2020). Motivation across a transition: Changes in achievement goal orientations and academic well-being from elementary to secondary school. *Learning and Individual Differences*, 79, 1-15. <https://doi.org/10.1016/j.lindif.2020.101854>

Vandenberg, R., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70. <https://doi.org/10.1177/109442810031002>

Wang, C., & Sun, T. (2020). Relationship between self-efficacy and language proficiency: A meta-analysis. *System*, 95, 102366. <https://doi.org/10.1016/j.system.2020.102366>

Yuen, M., & Chan, R. T. H. (2022). The influence of social connectedness and meaning in life on career adaptability and career self-efficacy in students with special educational needs. *International Journal for Educational and Vocational Guidance*. <https://doi.org/10.1007/s10775-022-09562-y>

Yuen, M., & Datu, J. A. D. (2021). Meaning in life, connectedness, academic self-efficacy, and personal self-efficacy: A winning combination. *School Psychology International*, 42(1), 79-99. <https://doi.org/10.1177/0143034320973370>

Zhou, J., Huebner, E., & Tian, L. (2020). Longitudinal Associations and Mechanisms Between Achievement Goals and Subjective Well-Being in School in Chinese Adolescents. *School Mental Health*, 12(2), 353-365. <https://doi.org/10.1007/s12310-019-09356-8>

Contact address: María de Lourdes Vargas Garduño. Universidad Michoacana de San Nicolás de Hidalgo, Facultad de Psicología. Francisco Villa, 450, Colonia Dr. Miguel Silva, Morelia, C.P. 58110, Michoacán, México. E-mail: maria.lourdes.vargas@umich.mx

A quantitative SWOT analysis for Spanish education

Un análisis DAFO cuantitativo para la educación española

<https://doi.org/10.4438/1988-592X-RE-2023-401-590>

Francisco López Rupérez

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

Universidad Camilo José Cela

Isabel García García

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

Universidad Camilo José Cela

Abstract

This paper aims to contribute to the strategic analysis of the Spanish system with the ultimate goal of helping to guide the improvement of its educational policy. In the field of education, SWOT analyses have so far focused on university institutions and, to a lesser extent, on schools. However, available SWOT analyses of national education systems are scarce and deficient. Two procedures have been used: a qualitative SWOT analysis (Strengths, Weaknesses, Opportunities and Threats) and the subsequent application of the quantitative technique of the Analytic Hierarchy Process (AHP). The preparation of the SWOT matrix has been carried out at two hierarchical levels: that of the sub-factors and that of the indicators. For internal factors (WS), resources, processes/policies (governance) and results have been taken as the basis. For external factors (TO), the PESTEL model has been adopted, limited to the political, socio-economic and technological categories. The specification of the sub-factors in indicators has been supported by the results of a significant number of national and international research and statistics. The subsequent application of the AHP technique has identified relative priorities based from comparisons within the multilevel hierarchical structures. The initial 38 SWOT indicators have been reduced to 17, of which, with a global priority of more than 0.050 and a low consistency ratio, the following stand out: intergenerational transmission of the educational level of parents, new technological tools for improving performance, Next Generation funds, lack of a

basic political agreement, and low level of excellence. Based on the results, the discussion focused on analysing the strategies of the *maxO-minW* and *minT-maxO* types and on formulating evidence-based recommendations aimed at implementing the resulting strategies.

Keywords: educational governance, Sstrategic planning of education, education policy, SWOT analysis, analytic hierarchy process.

Resumen

El presente trabajo pretende contribuir al análisis estratégico del sistema español con la finalidad última de servir de ayuda para orientar hacia la mejora su política educativa. En el ámbito educativo, los análisis DAFO se han centrado, hasta ahora, en las instituciones universitarias y, en menor medida, en los centros escolares. Pero son escasos y deficientes los análisis DAFO disponibles sobre sistemas educativos nacionales. Se ha recurrido a dos procedimientos concatenados: un análisis DAFO (Debilidades, Amenazas, Fortalezas y Oportunidades) de carácter cualitativo y la aplicación posterior de la técnica cuantitativa del *proceso de jerarquía analítica* (AHP). La elaboración de la matriz DAFO se ha efectuado en dos niveles jerárquicos: el de los subfactores y el de los indicadores. Para los factores internos (D, F), se ha tomado como base las categorías de recursos, procesos/políticas (gobernanza) y resultados. Para los factores externos (A, O), se ha adoptado el modelo PESTEL, limitado a las categorías de lo político, lo socioeconómico y lo tecnológico. La concreción de los subfactores en indicadores ha estado avalada por los resultados de un número apreciable de investigaciones y de estadísticas, tanto nacionales como internacionales. La aplicación subsiguiente de la técnica AHP ha determinado prioridades relativas a partir de comparaciones dentro de las estructuras jerárquicas multinivel. Se han reducido los 38 indicadores DAFO iniciales a 17, de los cuales, con una prioridad global mayor de 0,050 y una baja ratio de consistencia, destacan: transmisión intergeneracional del nivel educativo de los padres, nuevas herramientas tecnológicas para la mejora del rendimiento, fondos *Next Generation*, falta de un acuerdo político básico, y bajo nivel de excelencia. A partir de los resultados, la discusión se ha centrado en analizar las estrategias de los tipos *maxO-minD* y *minA-maxO* y en formular recomendaciones inspiradas en evidencias y destinadas a implementar las estrategias resultantes.

Palabras clave: administración de la educación, planificación educativa, política educativa, análisis DAFO, proceso de jerarquía analítica.

Introduction

Comparative analyses, based on international indicators, have revealed a worrying stagnation of the Spanish education system over the last two decades, in comparison with other neighbouring countries which, starting from similar or even inferior situations, have been able to react with determination and success (OECD, 2019; European Commission, 2020; López Rupérez & García García, 2020; OECD, 2021).

One of the facts that needs to be incorporated into the new global equation is that, as a result of an underlying complexity in which inter-dependencies proliferate and unexpected phenomena emerge (López Rupérez, 2021), education has become an indisputable part of the interactions between the global economy and society, and it is essential to manage it well in policy terms.

In this context, the use of so-called strategic thinking is a necessary condition for qualitative improvement. According to the Center for Management & Organization Effectiveness (2019):

Strategic thinking is simply an intentional and rational thought process that focuses on the analysis of critical factors and variables that will influence the long-term success of a business, a team or an individual (...) Strategic thinking requires research, analytical thinking, innovation, problem-solving skills, communication and leadership skills, and decisiveness (p. 1).

The aim of this paper is to contribute to the strategic analysis of the Spanish education system with the ultimate aim of helping to guide its educational policy towards improvement. To this end, two concatenated procedures have been used: a qualitative SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis and the subsequent application of the quantitative technique of the analytical hierarchy process (AHP).

Since its origins in the last century (Codina Jiménez, 2011; Santos-Caballero & Gil-LaSource, 2017; Benzaghta et al., 2021), the number of articles published on SWOT analysis has been accelerating, particularly since the beginning of this century (Santos-Caballero & Gil-LaSource, 2017). In education, the focus has been on university institutions and, to a lesser extent, on schools (Benzaghta et al., 2021). Khalid et al. (2017) have conducted a SWOT study on higher education in Pakistan, and

Velmonte (2020) has applied this technique to the education system in the Philippines, albeit with a very limited analytical focus. In Spain, the Autonomous Community of Castile and León has conducted a qualitative SWOT analysis of its education system as part of its *II Plan de atención a la diversidad en la educación* [II Care plan for Diversity in Education] (BOCYL, 2017). However, we have not found quantitative SWOT analyses of national education systems that are sufficiently rigorous and complete. For this reason, and despite the exploratory nature of our study, it opens up a way to introduce this type of analysis into the highly difficult task of strategic governance of education systems (López Rupérez & García García, 2022).

Methods

Applying the SWOT technique to the Spanish education system

Within what is understood as strategic management (Koontz et al., 2012), SWOT analyses are regarded, in the world of organizations, as a considerably useful means for the subsequent formulation of strategies and their eventual implementation. Figure I show the main conceptual elements of a SWOT analysis.

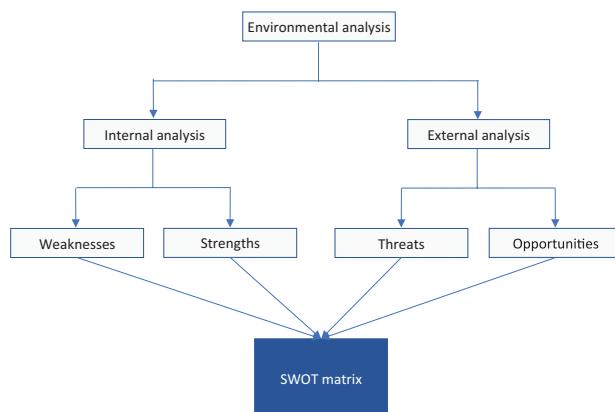
In the present study, the application to the Spanish education system of the procedural scheme represented in figure II has been carried out at two levels of concreteness in addition to that of the factors: that of the sub-factors and that of the indicators. As far as internal factors are concerned, their expression at the level of sub-factors has been based on the systemic approach, as is characteristic of the world of international education indicators (CERI-OECD, 1992), albeit centred on the three categories: resources, processes/policies (governance) and results. Regarding the external factors, we have based ourselves on the PESTEL model (Shilei & Yong, 2009; Yüksel, 2012; Jadan, 2020), but limiting this to the political, socio-economic, and technological categories. The general criterion for selecting the sub-factors in terms of indicators was relevance, which is supported by the results of a considerable amount of research and statistics, both national and international.

FIGURE I. Graphical representation of the typical factor matrix structure of a SWOT analysis



Source: Compiled by author based on linkedin.es

FIGURE II. Outline of the procedure for the elaboration of a SWOT matrix



Source: Compiled by author

Applying the analytical hierarchy process to the SWOT matrix

The *Analytic Hierarchy Process* (AHP) is a mathematical procedure created by Saaty (1980) and applicable in multi-criteria decision making, which allows a complicated problem to be broken down into a multi-level hierarchical structure of objectives, criteria, and alternatives (Sharma et al., 2008).

The transformation of the qualitative SWOT strategic analysis into a quantitative analysis using AHP (Kurttila et al., 2000) overcomes two limitations of the former: that it does not allow determining the relative importance of factors and sub-factors, and that it does not consider the combinations of the numerous criteria that can be considered, as well as their potential interdependencies (Pesonen, et al., 2001).

The absolute scales of factors and sub-factors are transformed by the AHP procedure into relative priorities based on comparisons within multilevel hierarchical structures (Saaty & Vargas, 1996). This is done based on a comparative scale –developed by Saaty to represent the relative importance of criteria, factors, or sub-factors– which is shown in Table I.

The pairwise comparison, whose elements represent alternatives, is arranged in a matrix that subsequently makes it possible to calculate the relative importance of the criteria, factors or sub-factors (Görener et al., 2012), and to determine the consistency ratios (Saaty & Vargas, 1996) that

TABLE I. AHP scale for a pairwise comparison of criteria, factors or attributes

Degree of importance	Description
1	Both criteria contribute equally to the objective.
3	Experience and judgement lean slightly in favour of one over the other.
5	Experience and judgement lean strongly in favour of one over the other.
7	Judgement is strongly favoured and its predominance is demonstrated in practice.
9	Extreme or absolute importance of one attribute over the other.
2, 4, 6 and 8	Used to represent trade-offs between the assessments described above.

Source: Saaty, 1980; Görener, Toker and Uluçay, 2012.

make it possible to assess the degree of consistency of the judgements with respect to that which would be provided by a large sample of purely random judgements¹.

Results

The resulting SWOT matrix for the Spanish education system

In accordance with the structure of factors and sub-factors described in the Methods Section, and based on empirical information derived from research and national and international statistics, the SWOT matrix described below has been developed.

■ Weaknesses

– *Inputs* (resources)

- WI1. Low public education expenditure relative to GDP (López Rupérez & García García, 2020; OECD, 2021; Montes-Pineda & López Rupérez, 2022).
- WI2. Low cumulative expenditure per pupil (6 years to 15 years) (OECD, 2021).
- WI3. Notable territorial inequality (López Rupérez et al., 2018a; 2018b)

– Governance (policies)

- WG1. Insufficient attention to evidence in the formulation and implementation of education policies (López Rupérez et al., 2017; López Rupérez et al., 2020 a; López Rupérez, 2022).
- WG2. Insufficient attention to accountability (López Rupérez et al., 2017).
- WG3. Deficient initial and in-service teacher training systems (López Rupérez et al., 2021).
- WG4. Deficient systems for access to school leadership and professional development (Leithwood et al., 2006; Hanushek et al., 2016; Pont Ferrer, 2017).

– Results (*outputs*)

- WR1. Underachievement in basic skills (European Commission, 2020).

¹ For a detailed description of the mathematical procedure and its calculation algorithms, see Coyle (2004)

- WR2. Low level of excellence (OECD, 2016; OECD, 2019).
- WR3. High early educational dropout (European Commission, 2020).

■ Strengths

- *Inputs* (resources)
 - SI1. Scholarships and study grants system² (OECD, 2021).
 - SI2. Level of teachers' salaries (Eurydice, 2021; OECD, 2021).
 - SI3. Educational expenditure in private institutions (OECD, 2021).
 - Governance (policies)
 - SG1. A consolidated educational bureaucracy.
 - SG2. Formal mechanisms for cooperation between education administrations³
 - SG3. A plural educational offer (Sainz & Sanz, 2021).
- *Results* (*outputs*)
 - SR1. High rates of early childhood education (European Commission, 2021).
 - SR2. High enrolment rates in primary and secondary education.
 - SR3. High rates of tertiary education graduates (ISCED 5-8) (European Commission, 2021).

■ Threats

- Political
 - TP1. The lack of a basic political agreement.
 - TP2. Weak political opposition on education.
 - TP3. The comparative advantage, in terms of education policy, of competing countries (Council of Europe, 2021; OECD, 2010; López Rupérez & García García, 2020).
- Socio-economic
 - TS1. A clear change in the economic and financial policy of the EU and the ECB.
 - TS2. The increase in social spending due to population ageing (INE, 2020).

² Statistics on Grants and Study Aids. General Subdirectorate of Statistics and Studies of the Ministry of Education and Vocational Training.

³ <https://www.educacionyfp.gob.es/mc/conferencia-sectorial-educacion/funcionamiento.html>

- TS3. The failure to integrate young people into the labour market^{4,5}
- TS4. The effects of pandemics (Alimi et al., 2021; Sanz et al., 2020).
 - Technological
 - TT1. A pressure on the system as an indirect effect of the reduction of jobs (Berggruen & Gardels, 2012; Baldwin, 2019).
 - TT2. Risk of depersonalization of the teacher-student relationship.
 - TT3. Interferences, due to improper use of technologies, with basic learning processes (OMS, 2022).
- Opportunities
 - Political
 - OP1. Expectations of political change (López Rupérez, 2021).
 - OP2. EU pressure on education (Consejo de la Unión Europea, 2002; Consejo Europeo, 2021).
 - OP3. The growing international information on successful educational policies (López Rupérez & García García, 2021).
 - Socioeconomic
 - OS1. The Next Generation EU funds⁶.
 - OS2. The demographic reduction in the number of pupils (INE, 2020).
 - OS3. Intergenerational transmission of parents' educational attainment (INE, 2019; OECD, 2021).
 - Technological
 - OT1. Remote and global interactions between education actors.
 - OT2. New technological tools for improving student performance (Patrick et al., 2013; Luckin & Issroff, 2018; López Rupérez, 2020).
 - OT3. Operational development of learning analytics (Ferguson et al., 2016).

⁴ <http://estadisticas.mecd.gob.es/EducaDynPx/educabase/index.htm?type=pcaxis&path=/laborales/insercion/afil&file=pcaxis&l=s0>

⁵ <https://www.universidades.gob.es/portal/site/universidades/menuitem.78fe777017742d34e0acc310026041a0/?vgnextoid=b747122d36680710VgnVCM1000001d04140aRCRD>

⁶ <https://www.lamoncloa.gob.es/temas/fondos-recuperacion>

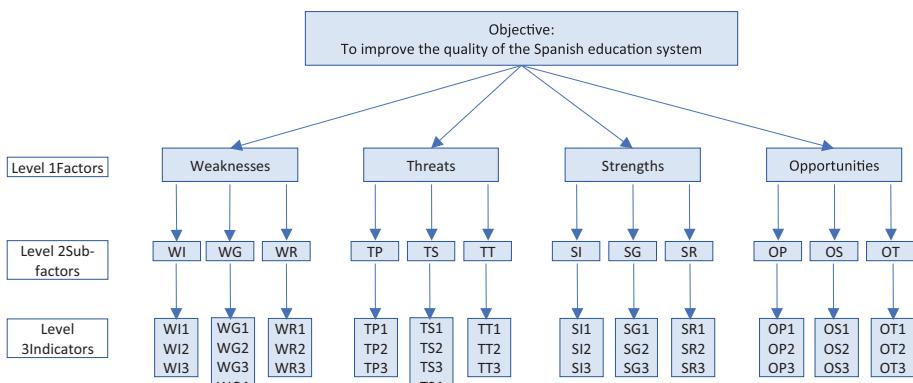
The result of the quantitative analyses

Applying the AHP procedure, with the determination of each matrix of weights — which have been independently assessed by the two authors, based on the pairwise comparisons carried out on the basic Saaty scale (table I) — has yielded the coinciding results shown in the tables in the annex, which refer to the three groups of variables corresponding to the respective hierarchical levels (see figure III).

The following three figures (IV, V and VI) show the results of the multilevel quantitative analysis. Thus, figure IV shows the SWOT factors and their sub-factors (levels 1 and 2), together with the figures representing the partial and global — or composite— priorities resulting from taking into consideration the different levels analysed. The highest relative values for each stage of analysis are highlighted in bold.

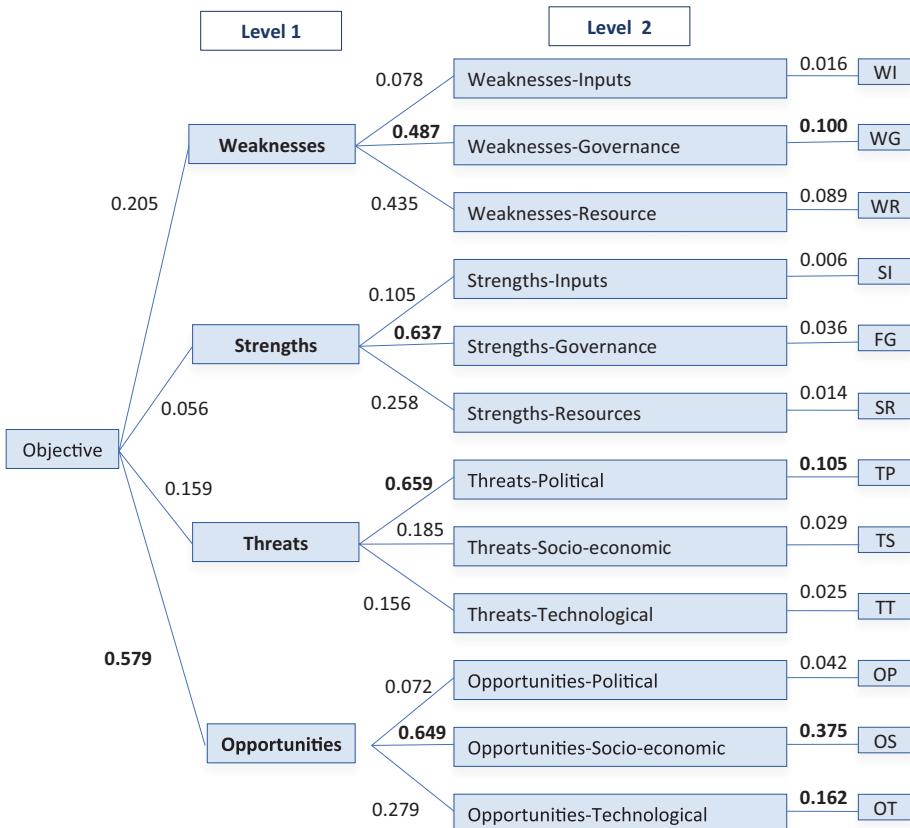
Figures V and VI show the sub-factors (level 2), their expression in indicators (level 3) as well as the corresponding figures for partial and global priorities, calculated in a similar way to that used in the previous phase. The complexity of the level 3 display required the use of two charts, one for the display of internal factors (W and S) and the other for external factors (T and O).

FIGURE III. Structure in hierarchical levels of the SWOT matrix



Source: Compiled by author

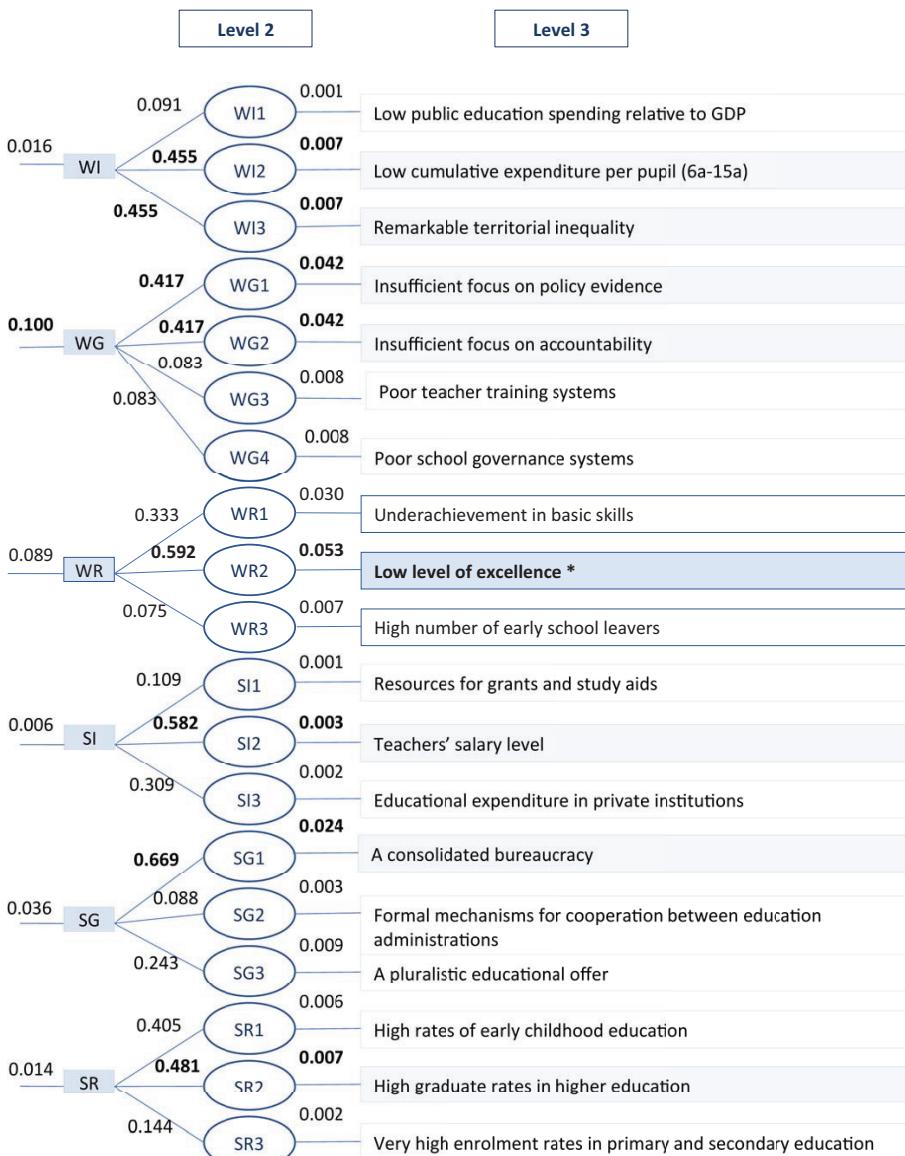
FIGURE IV. Overall scores for factors, sub-factors and SWOT indicators. Factors and sub-factors



Source: Compiled by author

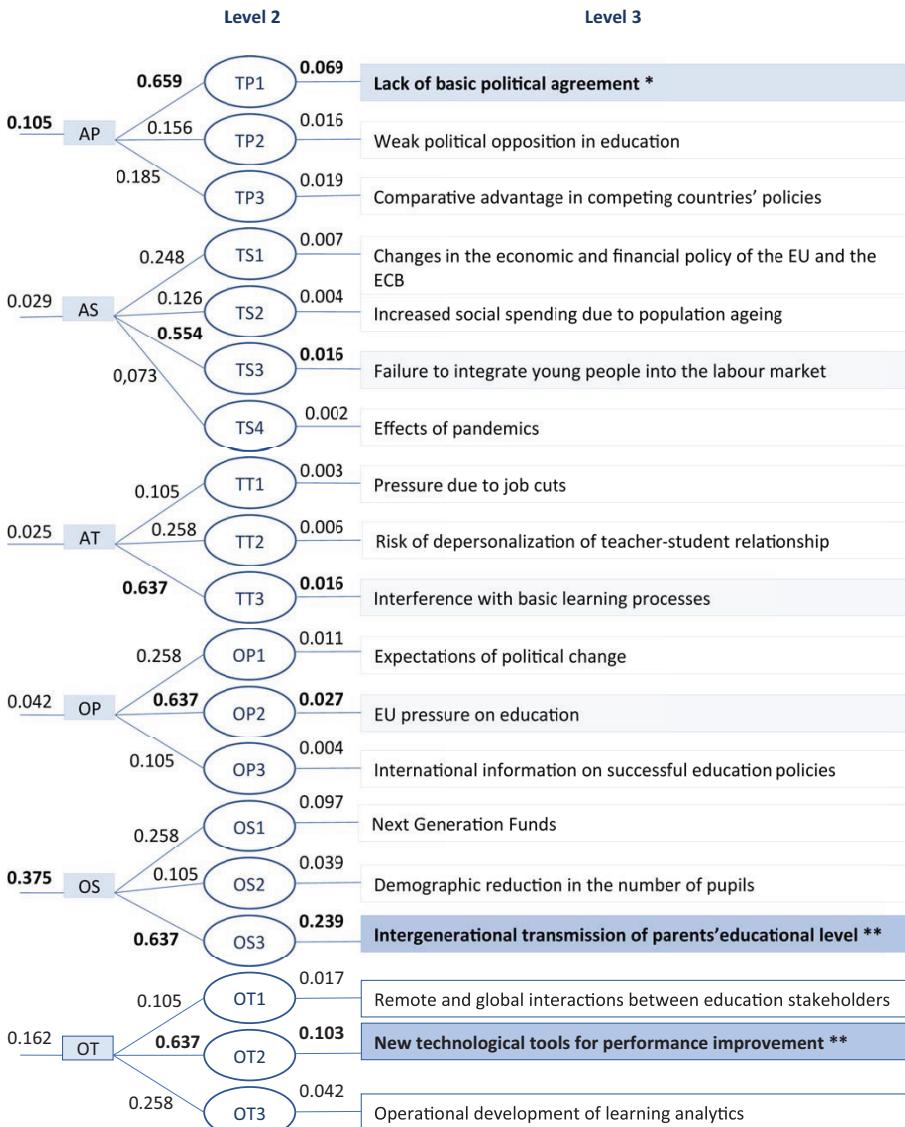
On the other hand, Table II provides a synthetic view of those indicators whose global priorities are equal to or higher than 0.010, together with their corresponding figures, as well as the CR consistency ratios, which will be useful in the discussion of strategic elements to be addressed later.

FIGURE V. Overall scores of factors, sub-factors and SWOT indicators. Internal factors, sub-factors and indicators



Source: Compiled by author

FIGURE VI. Overall scores for factors, sub-factors and SWOT indicators. External factors, sub-factors and indicators



Source: Compiled by author

TABLE II. Summary, in terms of priorities, of the quantitative analyses carried out on the SWOT matrix and its different hierarchical levels⁷

Factor	Level 1 Priority	Sub-factor	Indicator	Global priority
Weaknesses RC = 0.011	0.205	Governance RC = 0.000	Insufficient attention to evidence on policies	0.042
			Insufficient focus on accountability	0.042
		Results RC = 0.007	Low level of excellence	0.053 (*)
Strengths RC = 0.033	0.056	Governance RC = 0.006	A consolidated bureaucracy	0.024
Threats RC = 0.025	0.159	Political RC = 0.025	Lack of basic political agreement	0.069 (*)
			Weak political opposition in education	0.016
			Comparative advantage in political advantage of competitor countries	0.019
		Socio-economic RC = 0.073	Failure of young people to enter the labour market	0.016
		Technological RC = 0.033	Interferences in basic learning processes	0.016
Opportunities RC = 0.039	0.579	Political RC = 0.033	Expectations of political change	0.011
			EU pressure on education	0.027
			Next Generation EU funds	0.097 (*)
			Demographic decline in pupil numbers	0.039
		Socio-economic RC = 0.033	Intergenerational transmission of parental education level	0.239 (**)
			Remote and global interactions between educational actors	0.017
			New Technological tools for performance improvement	0.103 (**)
			Operational development of learning analytics	0.042

Source: Compiled by author

⁷ Note: (**) Global priority above 0.100; (*) Global priority between 0.050 and 0.100. Indicators with global priorities below 0.010 have been ignored in this Table.

Discussion

Applying the AHP technique allows SWOT analyses to be completed on a quantitative basis. These quantitative analyses aim to facilitate the adoption of strategic decisions informed by a multiplicity of criteria of different nature and relevance with varying degrees of interdependence. This SWOT enriched by AHP is a way of shedding light on a complex forest of relevant factors.

Table II highlights the prominent role of the Opportunities and provides a first reduction in complexity when suggesting strategies, since the systematic application of the AHP technique has reduced the initial thirty-eight relevant variables to less than half. Nevertheless, a series of heuristics –or guidelines based on experience– should be enunciated in order to advance in the problem of selecting relevant strategies resulting from the quantitative SWOT. What follows is a list, albeit not a full one, of those guidelines or orientations that we will apply:

- Start from a broad view of the available information (Gallego-Ayala & Juízo, 2011) (a).
- Focus on the most relevant pair of SWOT factors (Pesonen et al., 2001) (b).
- Rely mainly on offensive strategies (OF) and adaptive strategies (OD) (Weihrich, 1989; Codina Jiménez, 2011; Koontz et al., 2012) (c).
- Take into account the most important variables first (Gallego-Ayala & Juízo, 2011) (d).
- Assess consistency ratios together with the overall priority figures (Pesonen et al., 2001) (e).
- Test, on the basis of the study, the definition of an alternative strategy (Görenger et al., 2012) (f).

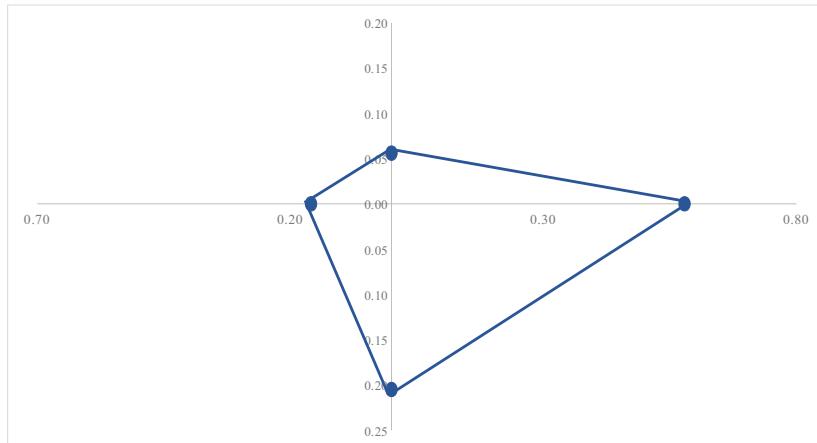
Based on the overview provided by Table II (orientation a), we now proceed to apply orientation b). Figure VII shows the figures for the priorities of level 1, the level corresponding to the SWOT factors. In accordance with the relative priority values of this first level, the strategies that revolve around opportunities, namely *maxO-maxS* (offensive strategy) and *maxO-minW* (adaptive strategy), are chosen (Koontz et al., 2012). The consideration of orientations c), d) and e) recommends focusing on the *maxO-minW* adaptive strategies, in accordance with the following approach:

- On the side of Opportunities:
 - Intergenerational transmission of parents' educational attainment.
 - New Technological tools for performance improvement.
 - Next Generation Funds.
- On the side of Weaknesses:
 - Low level of excellence.
 - Insufficient attention to evidence on policies.
 - Insufficient attention to accountability.

The inclusion of the selected governance indicators in the weaknesses group (see Table II) is justified, on the one hand, by the proximity of their priority scores to those of the results indicators and, on the other hand, by the advantage of this first group of indicators in terms of consistency (orientation e).

Finally, in view of the different scores in Table II, the application of orientation f) leads to an alternative and atypical strategic approach along the following lines: *minT-maxO*. In other words, minimizing the impact of the threats and maximizing the use of the opportunities, which means adding from the side of the Threats: Lack of basic political agreement.

FIGURE VII. Graphical representation of the priority scores obtained for the different level 1 SWOT factors.



Source: Compiled by author

It is now a matter of complementing this approach towards *what* should be done with an approach towards *how* to achieve it, and of formulating recommendations, as evidence-based as possible, to the public authorities, which help to implement the strategies that emerge from the orderly application of the heuristics.

Maximizing the use of intergenerational transmission

It is clear that all internal factors that contribute to this educational transmission from families will contribute to seizing this opportunity. According to the available empirical evidence, improving the quality of teachers and improving the quality of school leadership are the two most critical factors (Hattie, 2003; Leithwood et al., 2006; Hanushek et al., 2016; López Rupérez, 2021) for school success. In addition, a third factor, which lies at the heart of the process of intergenerational cultural transmission itself, is parental involvement (Castro et al., 2015), which is facilitated by successful cooperation between family and school.

How to take advantage of new technological tools for performance improvement

Two orientations would make it possible to take advantage of this opportunity. The first is to use digital technologies for the effective development of personalized –or learner-centred– teaching that reaches all students and addresses their needs and the particular characteristics of their learning process (López Rupérez, 2020). The second, closely related to the previous one, is to take advantage of what we know about the effectiveness of Mastery Learning with its precise and rigorous teaching sequences that ensure that all students master what they learn (López López, 2006).

How to optimize the use of Next Generation Funds in the educational field

If a Paretian approach to priority setting is adopted, then teacher-centred and school leadership-centred policies should be inexcusably part of the

objectives of *Lever VII. Education and knowledge, lifelong learning, and capacity development*.⁸

How to raise standards of excellence

Two empirically grounded recommendations can be made to achieve this goal: the first is to raise the level of teaching demands; the second is to strengthen non-cognitive skills, particularly those related to perseverance, resilience and a sense of effort (López Rupérez & García, 2017). It is clear that both recommendations are interlinked, because while raising the level of teaching demands must go hand in hand with raising the level of teacher competence, this will be insufficient if students are not simultaneously encouraged to develop these skills, which are generally associated with character development (Lickona & Davidson, 2005; Bernal et al., 2015).

How to enhance the role of knowledge and evidence in education policies

Two recommendations emerge from the decisive role of knowledge and evidence in defining policies and educational reforms in high-performing countries (López Rupérez, 2022). Firstly, an epistemological shift towards a critical rationalism that considers the principle of reality and respect for facts. It is a matter of introducing educational policies into the territory of rational-scientific approaches.

The second recommendation is of an instrumental nature, and consists of recovering the practice of 'white papers', which has been abandoned in Spain since the last century. No educational reform should be able to be included in the Official State Gazette without the prior drafting of a white paper which, together with the objectives, would provide a justification for the reforms based on facts and not on mere rhetorical formulations; an effort should be made to explain –at least by means of plausible conceptual models with some empirical basis and, if possible, by means of causal models – the mechanisms through which the reform is expected to achieve

⁸ https://www.lamoncloa.gob.es/temas/fondos_recuperacion

its objectives; and an impact assessment plan should also be included in order to check whether or not these forecasts are being fulfilled.

How to strengthen accountability

Two recommendations for strengthening accountability are essential. The first is to include the overall assessment model of the education system among the elements of a basic political consensus that blocks piecemeal changes. This will guarantee the stability of the model, facilitate the comparability of time series and ensure that useful and minimally reliable inferences can be made.

The second recommendation is to make the institution responsible for the evaluation of the education system independent of the government, with high academic prestige and technical solvency, and accountable to Parliament. This is what the Portuguese government did with the creation of an Institute for Educational Evaluation as an autonomous and independent body (Crato, 2020).

How to minimize the impact of the lack of a basic political agreement

From the analysis of experience, two recommendations arise and are justified below. The first is to introduce the aforementioned rationality in the formulation of the policies. This essential attribute makes it possible to get it right; but it also makes the stability of educational reforms more likely, which is a necessary condition for their success. The second is to progress towards a social pact rather than a political pact. Social expectations in Spain regarding the need to articulate an educational pact are in the majority and, therefore, the obstacles may well be of a lower calibre than in the strictly political sphere.

Both strategies are interrelated. Spanish society as a whole is more sensitive to rational arguments than its political class and, of course, much less sensitive to those that respond to a logic of power. The fact that the social agreement comes before the political pact will be a stimulus for the latter not to become disengaged and will generate a certain opportunity for its materialization.

Limitations of the study

The margin of subjectivity inherent in the method used in this study could be reduced by increasing the number of experts involved either in the selection processes of sub-factors and indicators, or in the processes of assigning weights by pairs, or in both. The application of the Delphi procedure of expert consultation (Landeta, 1999) could be one such possibility for methodological consolidation.

Bibliographical references

Alimi, Y. D., Bernstein, A., Epstein, J., Espinal, M., Kakkar, M., Kochervar, D., & Werneck, G. (2021). Report of the Scientific Task Force on Preventing Pandemics. *Harvard Global Health Institute*. <https://cdn1.sph.harvard.edu/wp-content/uploads/sites/2343/2021/08/Preventing%20PandemicsAug2021.pdf>

Baldwin, D. (2019). *La convulsión globótica. Globalización, robótica y el futuro del trabajo*. Antoni Bosch editor.

Benzaghta, M. A., Elwälada, A., Mousa, M. M., Erkan, I. & Rahman, M. (2021). SWOT analysis applications: An integrative literature review. *Journal of Global Business Insights*, 6(1), pp. 55-73. <https://doi.org/10.5038/2640-6489.6.1.1148>

Berggruen, N., & Gardels, N. (2012). *Governance inteligente para el siglo XXI. Una vía intermedia entre occidente y oriente*. Taurus.

Bernal, A., González Torres, M. C., & Naval, C. (2015). La educación del carácter. Perspectivas internacionales. *Participación educativa*. Segunda Época/Vol.4, n.º 6, pp. 35-45.

BOCYL (2017). II Plan de Atención a la Diversidad en la Educación de Castilla y León 2017-2022. *Boletín Oficial de Castilla y León*, núm. 115, pp. 23109-23176.

Castro, M., Expósito-Casas, E., López-Martín, E.; Lizasoain, L.; Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational Research Review*, 14, pp. 33-46.

Center for Management & Organization Effectiveness (2019). *What is Strategic Thinking?* CMOE. <https://cmoe.com/>

CERI-OECD (1992). *The OECD International Education Indicators. A Framework for analyses*. OECD Publishing.

Codina Jiménez, A. (2011). Deficiencias en el uso del FODA. Causas y sugerencias. *Revista Ciencias Estratégicas*, vol. 19, núm. 25, enero-junio, pp. 89-100. <https://www.redalyc.org/articulo.oa?id=151322413006>

Consejo de la Unión Europea (2002). *Programa de trabajo detallado para el seguimiento de los objetivos concretos de los sistemas de educación y formación en Europa*. Documento 2002/C142/01.

Consejo Europeo (2021). *Resolución del Consejo relativa a un marco estratégico para la cooperación europea en el ámbito de la educación y la formación con miras al Espacio Europeo de Educación y más allá (2021-2030)*.

Coyle, G. (2004). The Analytic Hierarchy Process (AHP). En *Practical Strategy*. Open Access Material. AHP 1 © Pearson Education Limited.

Crato, N. (2020). Curriculum and Educational Reforms in Portugal: An Analysis on Why and How Students' Knowledge and Skills Improved. In *Audacious Education Purposes. How Governments Transform the Goals of Education Systems*. Reimers, F. M. (Ed.). Springer. <https://bit.ly/36frag>

European Commission (2020). *Monitor de la Educación y la Formación de 2020*. <https://education.ec.europa.eu/es/monitor-de-la-educacion-y-la-formacion>

Eurydice/European Commission/EACEA (2021). *Teachers' and School Heads' Salaries and Allowances in Europe – 2019/20. Eurydice Facts and Figures*. Publications Office of the European Union. https://eacea.ec.europa.eu/national-policies/eurydice/content/teacherssalaries-2021_en

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., & Vuorikari, R. (2016). Research Evidence on the Use of Learning Analytics: Implications for Education Policy. R. Vuorikari, J. Castaño Muñoz (Eds.). *Joint Research Centre Science for Policy Report*, EUR 28294 EN. <https://doi.org/10.2791/955210>

Gallego-Ayala, J., & Juízo, D. (2011). Strategic implementation of integrated water resources management in Mozambique: An A'WOT analysis. *Physics and Chemistry of the Earth*, 36, pp. 1103-1111.

Görenger, A., Toker, K., & Uluçay, K. (2012). Application of Combined SWOT and AHP: A Case Study for a Manufacturing Firm. *Procedia-Social and Behavioral Sciences* 58, pp. 1525-1534. <https://doi.org/10.1016/j.sbspro.2012.09.1139>. 8th International Strategic Management Conference.

Hanushek, E. A., Rivkin, S. G., & Schiman, J. C. (2016). Dynamic effects of teacher turnover on the quality of instruction. *Economics of Education Review*, 55, pp. 132-148.

Hattie, J. (2003). *Teachers Make a Difference: What is the research evidence?* Australian Council for Educational Research Annual Conference on: Building Teacher Quality. October 2003, pp. 1-17.

INE (2019). INEbase. Encuesta de condiciones de vida. Módulo año 2019. <https://www.ine.es/jaxi/Table.htm?path=/t25/p453/modulo/2019/10/&file=07001a.px&L=0>

INE (2020). INEbase. Proyecciones de población. Proyección de la población de España. 2020-2070. <https://www.ine.es/jaxiT3/Table.htm?t=36643&L=0>

Jadan, D. (2020). *PESTEL analysis as a Baseline to support decision-making in the local textile industry*. International Conference on Systems and Information Sciences, Manta, Ecuador.

Khalid, J. Alia, A.J., Khaleela, M., Islama, M.S., & Shua, Q.F. (2017). Towards Global Knowledge Society: A SWOT Analysis of Higher Education of Pakistan in Context of Internationalization. *Journal of Business*, 02(02), pp. 08-15. <http://dx.doi.org/10.18533/job.v2i2.65>

Koontz, H., Wehrich, H., & Cannice M. (2012). *Administración. Una perspectiva global y empresarial*. 14^a edición. McGrawHill.

Kurttila, M., Pesonen, M., Kangas, J., & Kajanus, M. (2000). Utilizing the analytic hierarchy process (AHP) in SWOT analysis—A hybrid method and its application to a forest-certification case. *Forest Policy and Economics*, 1(1), pp. 41-52.

Landeta, J. (1999). *El método Delphi. Una técnica de previsión para la incertidumbre*. Editorial Ariel.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). *Seven strong claims about successful school Leadership*. Nottingham, England: National College of School leadership. <http://dx.doi.org/10.1080/13632430701800060>

Lickona, T., & Davidson, M. (2005). *Smart & Good High Schools. Integrating excellence and ethics for success in school, work, and beyond*. Center for the 4th and 5th Rs (Respect & Responsibility). Character.

López López, E. (2006). El *mastery learning* a la luz de la investigación educativa., 340. Mayo-agosto 2006, pp. 625-665.

López Rupérez, F. (2020). El currículo y la educación en el siglo XXI. La preparación del futuro y el enfoque por competencias. Narcea Ediciones.

López Rupérez, F. (2021). *La gobernanza de los sistemas educativos. Fundamentos y orientaciones*. Narcea-Stamp UCJC.

López Rupérez, F. (2022). La calidad de la gobernanza del sistema educativo. El caso de la LOMLOE. *Revista Española de Pedagogía*, 80 (281), 55-74. <https://doi.org/10.22550/REP80-1-2022-02>.

López Rupérez, F., & García García, I. (2017). *Valores y éxito escolar. ¿Qué nos dice PISA 2015?* Universidad Camilo José Cela. https://www.ucjc.edu/wp-content/uploads/valoresyexito_171116.pdf

López Rupérez, F., & García García, I. (2020). España vs. Portugal en educación. Una aproximación sistémica. *Revista Iberoamericana de Educación*, 84 (1), pp.193-215. <https://rieoei.org/RIE/article/view/4042>

López Rupérez, F., & García García, I. (2021). *El profesorado de educación secundaria a la luz de PISA 2018. Implicaciones para la política educativa española*. Universidad Camilo José Cela. <http://repositorio.ucjc.edu/handle/20.500.12020/925>

López Rupérez, F., & García García, I. (2022). *Hacia una Gobernanza Estratégica. Un análisis DAFO cuantitativo para la educación española*. Universidad Camilo José Cela. [hacia_una_gobernanza_estrategica.pdf \(ucjc.edu\)](http://hdl.handle.net/20.500.12020/825)

López Rupérez, F., García García, I., & Expósito Casas, E. (2017). *La calidad de la gobernanza del sistema educativo español. Un estudio empírico*. Universidad Camilo José Cela. <http://hdl.handle.net/20.500.12020/825>

López Rupérez, F., García García, I., & Expósito Casas, E. (2018 a). *PISA 2015 y las Comunidades Autónomas españolas. Diagnósticos empíricos y Políticas de mejora*. Universidad Camilo José Cela. <http://hdl.handle.net/20.500.12020/826>

López Rupérez, F., García García, I. and Expósito Casas, E. (2018 b). Educational Effectiveness, Efficiency, and Equity in Spanish Regions: What Does PISA 2015 Reveal? *Orbis Scholae*, 2018, 12 (2), pp. 9-36. <https://doi.org/10.14712/23363177.2018.291>

López Rupérez, F., García García, I., & Expósito Casas, E. (2020). Un marco analítico para la evaluación de la calidad de la gobernanza de los sistemas educativos. *Revista Iberoamericana de Educación*, 83 (1), pp. 53-76. <https://doi.org/10.35362/rie831367>

López Rupérez, F., García García, I., & Expósito Casas, E. (2021). Formación inicial y formación permanente del profesorado de educación secundaria en España. Un análisis territorial. *Bordón. Revista de Pedagogía*, 73 (4), 2021, pp. 65-84. <https://doi.org/10.13042/Bordon.2021.91081>

Luckin, R., & Issroff, K. (2018). *Future of Education and Skills 2030: Conceptual Learning*. EDU/ EDPC(2018)45/ANN2. OECD. <https://www.oecd.org/education/2030-project/curriculum-analysis>

Montes-Pineda, O., & López Rupérez, F. (2022). Gasto y Political educativas en España. La relevancia de la asignación económica para el rendimiento académico. *EduPsykhé. Revista de Psicología y Educación*, Vol. 19(1), pp.1-34.

OECD (2010). *The high cost of low educational performance. The long-run economic impact of improving PISA outcomes*. OECD. <https://www.oecd.org/pisa/44417824.pdf>

OECD (2016). *PISA 2015. Results (Volume II): Policies and Practices for Successful Schools*. OECD Publishing. <https://doi.org/10.1787/9789264267510-en>

OECD (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing. <https://doi.org/10.1787/5f07c754-en>

OECD (2021). *Education at a Glance 2021: OECD Indicators*. OECD Publishing. <https://doi.org/10.1787/b35a14e5-en>

OMS (2022). CIE-11 para estadísticas de mortalidad y morbilidad. 6C51 Trastorno por uso de videojuegos. <https://icd.who.int/browse11/l-m/es>

Patrick, S., Kennedy, K., & Powell, A. (2013). *Mean What You Say: Defining and Integrating Personalized, Blended and Competency Education*. iNACOL. The International Association for K-12 Online Learning. <https://aurora-institute.org/resource/mean-what-you-say-defining-and-integrating-personalized-blended-and-competency-education>

Pesonen, M., Ahola, J., Kurtila M., Kajanus, M., & Kangas, J. (2001). Applying A'WOT to Forest Industry Investment Strategies: Case Study of a Finnish Company in North America. In: Schmoldt, D. L., Kangas, J., Mendoza, G. A., Pesonen, M. (eds) *The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making. Managing Forest Ecosystems*, vol 3. Springer. https://doi.org/10.1007/978-94-015-9799-9_12

Pont Ferrer, B. (2017). *Reformas educativas: el caso del liderazgo escolar en perspectiva comparada*. (Tesis doctoral). Universidad Complutense de Madrid.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill.

Saaty, T. L., & Vargas, L. G. (1996). *Decision Making with The Analytic Network Process*. Springer.

Sainz, J., & Sanz, I. (2021). Los centros públicos y concertados se refuerzan mutuamente. Revisión de la Literatura. *Cuadernos de Pedagogía*, nº 516, pp. 16-20.

Santos-Caballero, A., & Gil-LaSource, J. (2017). La nueva Fuzzy-SWOT: Aplicación empírica con expertones. *Esic Market Economics and Business Journal*, Vol. 48, N.º 1, Enero-Abril, pp. 143-164.

Sanz, I., Sainz, J & Capilla, A. (2020). *Efectos de la crisis del coronavirus en la educación. Informe*,OEI.

Sharma, M. J., Moon, I., & Bae, H. (2008). Analytic hierarchy process to assess and optimize distribution network. *Applied Mathematics and Computation*, Vol. 202, pp. 256-265.

Shilei, L., & Yong, W. (2009). Target-oriented obstacle analysis by PESTEL modeling of energy efficiency retrofit for existing residential buildings in China's northern heating region. *Energy Policy*, 37, pp. 2098-2101. <https://doi.org/10.1016/j.enpol.2008.11.039>

Velmonte, G. L. (2020). SWOT Analysis Philippine Educational System. *International Journal of Intelligent Computing and Technology* (IJICT), Vol. 4, Iss.1, pp. 18-24.

Weihrich, H. (1989). *Excelencia Administrativa. Productividad mediante administración por objetivos*. Mc Graw Hill México.

Yüksel, I. (2012). Developing a Multi-Criteria Decision Making Model for PESTEL Analysis. *International Journal of Business and Management*. Vol. 7, No. 24 pp. 52-66.

Contact address: Francisco López Rupérez. Universidad Camilo José Cela, Facultad de Educación, Cátedra de Políticas Educativas. Castillo de Alarcón, 49. Urb. Villafranca del Castillo, Villanueva de la Cañada, 28692. Madrid. Spain. E-mail: flopezr@ucjc.edu

Annex

Level 1. Pairwise comparison of SWOT factors

TABLE A.1. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the SWOT factor group

SWOT Group	W	S	T	O	Degree of importance
Weaknesses (W)	1	5	1	1/3	0.205
Strengths (S)	1/5	1	1/3	1/7	0.056
Threats (T)	1	3	1	1/5	0.159
Opportunities (O)	3	7	5	1	0.579
RC = 0.039					

Source: Compiled by author

Level 2. Pairwise comparison of the groups of sub-factors in which each SWOT factor is expressed

TABLE A.2. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the subgroup for the Weakness (W) factor.

WEAKNESSES Group	I	G	R	Degree of importance
Inputs (I)	1	1/7	1/5	0.078
Governance (G)	7	1	1	0.487
Results (A)	5	1	1	0.435
RC = 0.011				

Source: Compiled by author

TABLE A.3. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the subgroup corresponding to the Strength (S) factor.

STRENGTHS Group	I	G	R	Degree of importance
Inputs (I)	1	1/5	1/3	0.105
Governance (G)	5	1	3	0.637
Results (A)	3	1/3	1	0.258
RC = 0.033				

Source: Compiled by author

TABLE A.4. Pairwise comparison matrix, degrees of importance and Consistency Ratio for the Threats (T) factor subgroup.

THREATS Group	P	S	T	Degree of importance
Political (P)	1	3	5	0.659
Socio-economic (S)	1/3	1	1	0.185
Technological (T)	1/5	1	1	0.156
RC = 0.025				

Source: Compiled by author

Level 3. Pairwise comparison of the groups of indicators in which each of the sub-factors – Inputs, Governance and Results, and Political, Socio-economic, and Technological – are expressed

TABLE A.5. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators corresponding to the Weaknesses-Inputs (WI) sub-factor.

WEAKNESSES-INPUTS Subgroup	WI1	WI2	WI3	Degree of importance
Low public education expenditure relative to GDP (WI1)	1	1/5	1/5	0.091
Low cumulative expenditure per pupil (WI2)	5	1	1	0.455
Notable territorial inequality (WI3)	5	1	1	0.455
RC = 0.000				

Source: Compiled by author

TABLE A.6. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators for the Weaknesses-Governance (WG) sub-factor

WEAKNESSES-GOVERNANCE Subgroup	DG1	DG2	DG3	DG4	Degree of importance
Insufficient attention to evidence on policies (WG1)	1	1	5	5	0.417
Insufficient attention to accountability (WG2)	1	1	5	5	0.417
Weak teacher education systems (WG3)	1/5	1/5	1	1	0.083
Weak school leadership systems (WG4)	1/5	1/5	1	1	0.083
RC = 0.000					

Source: Compiled by author

TABLE A.7. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators corresponding to the Weaknesses-Results (WR) sub-factor.

WEAKNESSES-RESULTS Subgroup	WR1	WR2	WR3	Degree of importance
Low performance in basic skills (WR1)	1	1/2	5	0.333
Low level of excellence (WR2)	2	1	7	0.592
High level of early school leavers (WR3)	1/5	1/7	1	0.075
RC = 0.007				

Source: Compiled by author.

Note: The clear preponderance in the assessment of indicator WR2 over WR3 is mainly justified by the different nature of the underlying data source: the former are derived from objective evidence, the latter are of purely administrative origin and therefore modifiable.

TABLE A.8. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the indicator set for the Strengths-Inputs (SI) sub-factor.

STRENGTHS-INPUTS Subgroup	SI1	SI2	SI3	Degree of importance
Resources for Scholarships and Grants (SI1)	1	1/5	1/3	0.109
Level of teachers' salaries (SI2)	5	1	2	0.582
Educational expenditure on private institutions (SI3)	3	1/2	1	0.309
RC = 0.003				

Source: Compiled by author

Note: The lower relative strength of the SI1 indicator in the allocation of weights is a reflection of the evidence in the international comparison.

TABLE A.9. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the indicator set for the Strengths-Governance (SG) sub-factor.

STRENGTHS-GOVERNANCE Subgroup	SG1	SG2	SG3	Degree of importance
A consolidated bureaucracy (SG1)	1	7	3	0.669
Formal mechanisms of cooperation between educational administrations (SG2)	1/7	1	1/3	0.088
A plural educational offer (SG3)	1/3	3	1	0.243
RC = 0.006				

Source: Compiled by author

TABLE A.10. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the indicator set for the Strengths-Results (SR) sub-factor.

STRENGTHS-RESULTS Subgroup	SR1	SR2	SR3	Degree of importance
High rates of early childhood education (FR1)	1	1	3	0.405
High rates of tertiary education graduates (FR2)	1	1	5	0.481
Very high enrolment rates in primary and ESO (FR3)	1/3	1/5	1	0.114
RC = 0.025				

Source: Compiled by author

Note: The lower relative weight of the SR3 indicator is justified because, although it represents an intrinsic strength of the system, in comparative terms it is widespread in developed countries.

TABLE A.11. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators for the Threats-Political (TP) sub-factor.

THREATS-POLITICAL Subgroup	TP1	TP2	TP3	Degree of importance
Lack of basic political agreement (TP1)	1	5	3	0.659
Weak political opposition in education (TP2)	1/5	1	1	0.156
Comparative advantage in political advantage of competing countries (TP3)	1/3	1	1	0.185
RC = 0.025				

Source: Compiled by author

Note: Evidence shows how often the lack of political agreement leads to instability in education reforms in Spain.

TABLE A.12. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators for the SocioEconomic Threats (SE) sub-factor.

THREATS-SOCIOECONOMIC Subgroup	TS1	TS2	TS3	TS4	Degree of importance
Changes in the economic and financial policy of the EU and the ECB (TS1)	1	3	1/3	3	0.248
Increased social spending due to ageing of the population (TS2)	1/3	1	1/5	3	0.126
Failure of young people to enter the labour market (TS3)	3	5	1	5	0.554
Effects of pandemics (TS4)	1/3	1/3	1/5	1	0.073
RC = 0.073					

Source: Compiled by author

TABLE A.13. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators for the Threats-Technological (TT) sub-factor.

THREATS-TECHNOLOGICAL Subgroup	TT1	TT2	TT3	Degree of importance
Pressure due to job cuts (TT1)	1	1/3	1/5	0.105
Risk of depersonalization of the teacher-student relationship (TT2)	3	1	1/3	0.258
Interferences with basic learning processes (TT3)	5	3	1	0.637
RC = 0.033				

Source: Compiled by author

Note: The attribution of pairwise weights has taken into account the indirect nature of the effects of TT1 versus the direct and extensive nature of TT3.

TABLE A.14. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators corresponding to the Opportunities-Political (OP) sub-factor.

OPPORTUNITIES-POLITICAL Subgroup	OP1	OP2	OP3	Degree of importance
Expectations of policy change (OP1)	1	1/3	3	0.258
EU pressure on education (OP2)	3	1	5	0.637
International information Successful education policies (OP3)	1/3	1/5	1	0.105
RC = 0.033				

Source: Compiled by author

Note: The attribution of peer weights has taken into account the increasing EU pressure on education as a consequence of the parallel importance given to education and training, according to the policy approaches of the European Council.

TABLE A.15. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators for the Socio-Economic Opportunities (OS) sub-factor.

OPPORTUNITIES- SOCIOECONOMIC Subgroup	OS1	OS2	OS3	Degree of importance
Next Generation Funds (OS1)	1	3	1/3	0.258
Demographic reduction of pupil numbers (OS2)	1/3	1	1/5	0.105
Intergenerational transmission of parents' educational level (OS3)	3	5	1	0.637
RC = 0.033				

Source: Compiled by author

Note: The attribution of pairwise weights has taken into consideration the robust predictability of the OS3 indicator.

TABLE A.16. Pairwise comparison matrix, degrees of importance and Consistency Ratio of the set of indicators corresponding to the Opportunities-Technological (OT) sub-factor.

OPPORTUNITIES-TECHNOLOGICAL Subgroup	OT1	OT2	OT3	Degree of importance
Remote and global interactions between educational actors (OT1)	1	1/5	1/3	0.105
New Technological tools for performance improvement (OT2)	5	1	3	0.637
Operational development of learning analytics (OT3)	3	1/3	1	0.258
RC = 0.033				

Source: Compiled by author

Factors influencing the development of reading competencies in Peruvian students in second-grade of secondary education

Factores que influyen en el desarrollo de las competencias lectoras en estudiantes peruanos de 2º grado de secundaria

<https://doi.org/10.4438/1988-592X-RE-2023-401-591>

Manuel Alfredo Marcos Balabarcá

<https://orcid.org/0000-0001-6252-035X>

Universidad Complutense de Madrid

Abstract

This research focuses on identifying how various individual, family and school factors are related to the development of reading literacy among secondary school students in public schools in urban Peru. This competence is operationalised as the score obtained in the 2019 Censal Student Assessment. Among the most important results, it stands out that (for) in this specific context, pedagogical practices are not related to the development of reading literacy. It is the factors of the family context, specifically the socio-economic context, that have the greatest influence on this variable. This sets up an important problem, given that the evidence on school effectiveness and factors associated with learning indicate that the teaching action is the factor that should have the greatest influence on the development of student learning. Thus, in the last section of this document, some aspects that could be intervening and distorting the possible relationship between pedagogical practices and student learning are discussed.

Keywords: Learning, teacher, pedagogical practices, socioeconomic status, secondary school.

Resumen

La presente investigación se centra en identificar la forma cómo diversos factores individuales, familiares y escolares se relacionan con el desarrollo de la competencia lectora que presentan los estudiantes de secundaria de escuelas públicas del área urbana de Perú. Dicha competencia se operativiza como el puntaje obtenido en la Evaluación Censal de Estudiantes del año 2019. Entre los hallazgos más importante, se tiene que, las prácticas pedagógicas no guardan relación con el desarrollo de la competencia lectora. Por el contrario, son los factores del contexto familiar, en específico el socioeconómico, los que influyen en mayor medida sobre dicha variable. Esto configura una problemática importante, dado que la evidencia sobre la eficacia escolar y los factores asociados al aprendizaje señalan que la acción docente es el factor que debería tener una mayor influencia sobre el desarrollo de aprendizaje de los estudiantes. Así, en la última sección de este documento se discute sobre algunos posibles aspectos que podrían estar interviniendo y distorsionando la relación entre las prácticas pedagógicas y el aprendizaje de los estudiantes. Es necesario considerar esos factores para poder comprender y abordar la problemática mencionada previamente. La investigación presentada es un primer paso para entender la complejidad de la relación entre los diferentes factores y el desarrollo de la competencia lectora en los estudiantes de secundaria de Perú.

Palabras clave: Aprendizajes, docente, práctica pedagógica, nivel socioeconómico, escuela secundaria.

Introduction

The present research focuses on determining which factors would be related to the development of reading literacy achieved by students in second-grade of secondary education (ISCED 2) from public schools in urban areas of Peru. The results of the Student Census Evaluation (ECE) and the data provided by the different actors of the educational community (students, teachers and principals) were used in the questionnaires of associated factors collected together with said test.

A black box called learning

Academic performance reflects the development of competence in the education system and is measured by the average grade obtained by stu-

dents. This qualification must reflect specific learning and the ability to apply it in real situations (Ertmet & Newby, 2013), which is synthesized from the teacher's actions, their interaction with the socio-school context and the attitudes and aptitudes of the students. Thus, the teacher's way of teaching is the key to the development of student learning and competencies (Ausbel, 1976; Hammonds & Lamar, 1972; Lloyd & Fernyhough, 1999). The relationship between the teacher's actions, the school context and the socio-family and personal characteristics of the student influence the development of competencies (Coleman, 1966; La Serna, 2011, Martín et al., 2008). In addition, the socio-family characteristics of the student are important in explaining the differences in academic performance throughout schooling (Berger & Toma, 1994; Coleman, 1966; Hanushek and Taylor, 1990; Summer & Wolfe, 1977), especially in secondary school. (Martín et al., 2008; Risso et al., 2010).

The evidence on the characteristics of teachers is ambiguous. According to Berger and Toma (1994), there is no significant relationship between student performance and the teacher's years of work experience, nor between variables such as teacher's salary and educational level. However, Fuller and Clarke (1990), in their research for developing countries, found that the characteristics of schools play an important role in the development of competencies.

In line with the aforementioned (above), several studies have concluded that the school environment and socioeconomic conditions have an impact on the academic performance of students. According to Harbinson and Hanushek (1992), the quality of the school infrastructure and the learning materials used by the teacher are positively correlated with student performance in Brazil. In addition, access to basic services and technology is also related to improvements in academic performance. Mizala and Romaguera (1999) in Bolivia, found that socio-family and racial variables have a high degree of explanation of academic performance, followed by school management and teacher characteristics, especially experience and the use of homework.

Research in Peru reveals that pedagogical strategies and teacher training are important in explaining differences in performance between students. According to Beltrán and Seinfeld (2012) the educational level of the teacher influences performance in Reading and Mathematics. Cueto et al. (2008) find that curricular coverage and feedback strategies in class are positively related to performance in Reading and Mathematics, but

are correlated with students' socioeconomic status. Likewise, Servan and Tantalean (2011) found that teachers with postgraduate degrees have students with high scores in Reading and Mathematics. In addition, the use of didactic strategies and feedback are positively and importantly related to performance.

Some research indicates that the socioeconomic context of the student and the classroom, rurality, and the leadership of the school principal influence the effectiveness of the teacher and his practice for the development of competencies. According to Belfi et al. (2015) and Creemers and Kyriakides (2007), teachers in schools with students from vulnerable contexts are perceived as less effective than their peers in schools with students from more privileged contexts, since poverty hinders aspects such as planning, learning monitoring and school management. These findings suggest that student performance varies by context, both at the macro level between developed and developing countries, and for particular social strata and groups of individuals. In the Latin American context, the pedagogical leadership of principals is important in improving educational quality, especially in contexts of socioeconomic vulnerability (Freire & Miranda, 2014; Horn & Marfán, 2010; Leithwood et al., 2008, Ministry of Education, 2018).

The reviewed research indicates that the interaction between pedagogical practices, pedagogical leadership and student learning is crucial in the development of competencies. Thus, it is necessary to identify this relationship in specific contexts to develop tools and actions that improve the learning process and strengthen teaching action.

Good pedagogical practices and learning development

Constructivism refers to learning as an active construction by the student, adaptable to new knowledge (Phillips, 1995; Ortega & Romero, 2020; Von Glaserfeld, 1996). Thus, the National Curriculum for Basic Education (CNEB) of Peru adopts a competency-based approach within a socio-constructivist framework (Ministry of Education, 2016a). In addition, learning within schools is circumscribed to a set of activities directed towards a particular purpose, which is known as the teaching-learning process (La Serna, 2011). In this process, constructivism

positions the teacher as a mediator between the student and learning to develop useful skills for proper development in society (Hernández et al., 2003).

Considering empirical evidence, it has been identified that certain teacher pedagogical practices are positively related to the development of competencies, such as the appropriate climate in the classroom, mastery of the subject, exposure to learning opportunities, and collaboration between teachers. (Arcia & Laguna, 2004; Arenas and Ámbros, 2021; Fuller & Clarke, 1990; Martín et al., 2022; Monge & Gómez, 2020; Murillo & Martínez, 2018; Willms & Somer, 2001). Likewise, the teacher's decisions about the use of resources, strategies and types of materials to which students are exposed in the classroom (Balarin, 2016; Campos et al., 2014; Caro et al., 2015; Cueto et al., 2008; Hanushek & Woessmann, 2009).

Research in the Peruvian context has found a positive relationship between student achievement and collaboration between school teachers. Thus, Caro (2003) finds that meeting activities between teachers within the school is positively and significantly associated with the performance achieved by students in Mathematics. On the other hand, the Ministry of Education (2017) highlights the importance of collaborative work between teachers to enhance the impact of strategies for teaching reading (those focused on the meaning of the text and not on the repetition of content) and the improvement of pedagogical practice (Caro, 2003; Cueto et al., 2008). In addition, practices focused on repetition of content and passive resolution of exercises are negatively related to the development of competencies (Caro, 2003; Cueto et al, 2008).

According to the Ministry of Education (2016b) and Martínez et al. (2022), the impact of the teacher's teaching strategies on the development of reading is significant. Students with teachers who focus their strategies on content repetition obtain 9.0% of one standard deviation (SD) less performance than those with teachers with less passive strategies. In addition, in regions of socioeconomic vulnerability, such as the jungle, the coefficient increases to 11.3% of a SD. Likewise, the Ministry of Education (2017) mentions that students who receive a pedagogical practice that reflects a solid knowledge of the curricular area (mainly in Mathematics) and an approach with constructivist principles have an increase in their performance throughout schooling.

When strategies focused on the meaning of texts rather than repetition are used, the achievement gap between students of high and low socioeconomic status is reduced.

In summary, the pedagogical action of the teacher influences the development of competencies, and there are more effective pedagogical practices than others depending on the context in which they are applied. For example, the implementation of technologies may not have an impact on learning in vulnerable contexts. However, this relation is not absolute, it depends on how other factors influence the context. It is important to identify the relationships between pedagogical practices and learning in each particular context.

Reading literacy in the Student Census Assessment

The ECE is the main tool of the Peruvian Ministry of Education to measure the development of competencies at the national level. This large-scale standardized test aims to provide information on student learning achievements and compare the results over time to report on the evolution of learning (Ministry of Education, 2016b).

For the year 2019, the Ministry of Education (2020a) reports that students in 2nd and 4th grade of primary school were evaluated (165,658 and 125,540 students from 5,976 and 4,799 schools respectively) and 2nd grade of secondary school (511,874 students from 13,437 schools), using multiple-choice questions in a paper-and-pencil format. In addition, information was collected from different actors in the educational community (students, parents, teachers and principals) through questionnaires. For the secondary level, the Communication area test was built considering evaluating one of the competencies proposed by the CNEB. Specifically, the competence “Read various texts written in their mother tongue” and the three capacities defined for it: obtain information from the written text, infer and interpret the meaning of the text, and reflect on the form, content, and context of the text (Ministry of Education, 2016a, 2022a). The Communication test approached these capacities through five types of written texts, grouped into two textual formats and in three different types of contexts. Table I shows the organization of the texts that the evaluation had in this area.

Table I. Organization of the test content in the area of Communication in ECE

Types of text	Text format	Contexts
- Narrative	- Continuous	- Educational
- Expository	- Discontinuous	- Public
- Descriptive		- Recreational
- Argumentative		
- Instructive		

Source: Ministry of Education and Vocational Training (2020a)

Collection of teacher's pedagogical practices in ECE

The Office of Learning Quality Measurement (UMC) of the Minedu developed a questionnaire to collect information on the characteristics, perceptions and pedagogical practices of teachers, focusing on characteristics of the constructivist approach and seeking those that influence educational performance. The questionnaire included scales related to the teacher's pedagogical practice, including those related to the type of reading material the teacher uses and the frequency with which it is presented to students. Also, those related to the pedagogical practices of reading, classroom climate and students' perception of the pedagogical practice of their teachers (Rodríguez, Jenaro & Castaño, 2022). Table II shows the items of these scales.

Taking into account the limitations due to the availability of data, this research tries to identify and quantify the relationship between the pedagogical practices of the Communication teacher and the performance of students in urban public schools in Peru, using ECE data from 2019. Thus, it is hypothesized that these practices have an impact on student performance in Reading. The following sections detail the analytical strategy used to make the corresponding estimates and the results derived from its application in order to verify the proposed hypothesis.

Table II. Description of the scales related to pedagogical practices and the items that compose them

Group	Scale	Items that make up the scale
Use of materials	Reading different types of text	<ul style="list-style-type: none"> - Argumentative (opinion pieces, posters, newspaper editorials, etc.). - Expository (scientific papers, monographs, etc.). - Descriptive (encyclopedic articles, descriptive infographics, etc.). - Literary narrative text (anecdotes, short stories, comics, chronicles, etc.). - Non-literary narratives (biographies, news, etc.). - Instructive (manuals, recommendations text, recipes, etc.).
	Reading different types of text formats	<ul style="list-style-type: none"> - Continuous (organized in paragraphs without tables or diagrams). - Discontinuous (infographics, graphs, charts). - Mixed (contains paragraphs along with graphics or charts). - Multiples (two or more texts on the same subject).
Pedagogical practices	Pedagogical reading practices	<ul style="list-style-type: none"> - Encourage students to express their opinion about the texts read. - help students to relate the stories they read with their lives. - show students how the information in the text is based on what they already know. - I ask questions that motivate students to actively participate. - I set clear goals for student learning. - I ask questions to verify if the students have understood what I have taught them. - At the beginning of the class, I present a brief summary of the previous class. - I tell the students what they have to learn.
	Classroom climate	<ul style="list-style-type: none"> - Many students do not listen to what I say. - There is noise and disorder. - have to wait a long time for the students to be silent. - Students cannot work well. - Students do not begin work until long after class begins.
Student perception	Student perception of their teachers' pedagogical practices	<ul style="list-style-type: none"> - At the beginning of the classes they explain to us what we are going to do and learn. - At the beginning of the class they make us remember what we did in the previous class. - They use examples from everyday life and current events to explain the topics. - At the beginning of the class they ask us what we know about the subject that we will see. - They ask us to argue our ideas. - They ask us questions to make sure we have understood the topic. - When reviewing our work or exams, they leave us notes explaining what we should improve on. - They notice when we make a mistake when answering a question and they explain it to us. - They give us recommendations on how to better learn what we are taught. - They explain to us what we will learn when they give us a class assignment.

Methodology

Participants

The target population corresponds to communication students and teachers in 2nd grade of secondary school in urban state schools in Peru who participated in the ECE in 2019. The total number of students in the sample used was 114,093. from whom there was information on their reading performance and factors associated with their family background. Likewise, 5,568 teachers and 1,609 principals (or schools) who answered questionnaires with information on their personal and contextual characteristics were considered.

Variables

The criterion variable is the score attained by the student on the ECE in Reading. The predictor variables correspond to information collected in the associated factors questionnaires applied together with the test. Sociodemographic and perception variables of students, teachers and principals are considered. The student variables are presented in Table III. The scales were built by the UMC of the Minedu and are valid and reliable (Ministry of Education, 2020b).

Table III. Description of variables related to student characteristics

Indicator	Calculation	Description
Student performance	Rasch Model	Measure of a student's ability in a competency evaluated at a given grade level of schooling. It is a scale centered at 500 points.
Student gender	Student's report	0 = Male 1 = Female
Student's native language	Student's report	0 = Spanish 1 = Original language
Assistance to pre-school education	Student's report	0 = Did not attend pre-school education 1 = Attended pre-school education

(Continued)

Table III. Description of variables related to student characteristics (Continued)

Indicator	Calculation	Description
Repetition	Student's report	0 = Did not repeat any school grade 1 = Repeated a school grade
Socioeconomic Index	Principal component analysis	Score that represents the economic potential of the family to cover and satisfy needs, as well as to increase access to personal and social development opportunities.
Beliefs about reading: Reading as a natural talent	Confirmatory factor analysis	Student beliefs that reflect an understanding of reading ability as natural or unchangeable.
Beliefs about reading: Reading as a learned competence	Confirmatory factor analysis	Student beliefs that reflect an understanding of reading as a skill that can be developed with effort and time.
Reading self-efficacy	Confirmatory factor analysis	Student's perception of their ability to develop reading activities.
Pedagogical practices (general)	Confirmatory factor analysis	Student perception about the actions carried out by the teacher in order to facilitate and guide their learning.
Coexistence: manifestations of physical, verbal and/or psychological violence by teachers	Confirmatory factor analysis	This scale includes the forms of aggression carried out by teachers that the student witnesses, both physically and verbally or psychologically. In this sense, "physical violence" is understood as any act that causes damage or pain, and "verbal violence" or "psychological" as emotional damage without any kind of physical contact.
Coexistence: relationship between students	Confirmatory factor analysis	This scale seeks to find out how the student feels about the relationship they have with their classmates, and with this, to identify if the interpersonal relationships that develop are positive or not.
Coexistence: sense of belonging	Confirmatory factor analysis	This scale seeks to know the level of student identification with their school.

Source: Ministry of Education (2016b) and Ministry of Education and Vocational Training (2022b)

The characteristics and perceptions of teachers and principals are shown in Table IV. As with the student scales, these were prepared by the Minedu UMC and are valid and reliable (Ministry of Education, 2020b).

Table IV. Description of variables related to the characteristics of the teacher and principal

Indicator	Calculation	Description
Teacher gender	Teacher's report	0 = Male 1 = Female
Teacher's educational level Qualifications	Teacher's report	0 = Full college education at most. 1 = Postgraduate studies at least (master's and/or doctorate)
Teaching experience	Teacher's report	0 = Novice (Equal to or less than 5 years of teaching experience) 1 = Experienced (5+ years of teaching experience)
Curricular coverage	Confirmatory factor analysis	Frequency with which the Communication teacher addresses different topics related to the area in class.
Reading different types of text	Confirmatory factor analysis	Frequency with which the Communication teacher suggested to his students the reading of different types of texts.
Reading different text formats	Confirmatory factor analysis	Frequency with which the Communication teacher suggested to his students the reading of different text formats.
Self-efficacy on teaching reading	Confirmatory factor analysis	Confidence of the Communication teacher in his ability to develop reading competence among his students.
Reading as a natural talent	Confirmatory factor analysis	Communication teacher's belief that reading competence is an innate ability.
Reading as a learned competence	Confirmatory factor analysis	Communication teacher's belief that reading competence is a skill acquired with effort and dedication.
Reading self-efficacy	Confirmatory factor analysis	Confidence of the Communication teacher in his own reading competence.
Reading pedagogical practices	Confirmatory factor analysis	Frequency with which the Communication teacher implements a set of favorable practices to the learning of their students.
Interaction with other faculty	Confirmatory factor analysis	Frequency with which the Communication teacher interacts with their peers in order to improve the teaching process.
Classroom discipline climate	Confirmatory factor analysis	Frequency with which disruptive events occur in the class (noise, disorder, etc.) that arise from the teaching-learning processes.
Principal gender	Principal's report	0 = Male 1 = Female

(Continued)

Table IV. Description of variables related to the characteristics of the teacher and principal
(Continued)

Indicator	Calculation	Description
Principal's educational level	Principal's report	0 = Full college education at most. 1 = Postgraduate studies at least (master's and/or doctorate)
Principal experience	Principal's report	0 = Novice (Equal to or less than 5 years of teaching experience) 1 = Experienced (5+ years of teaching experience)
Pedagogical Leadership: Strategic Planning	Confirmatory factor analysis	Priority that the principal gives to activities related to the monitoring of planning, supervision, and achievement of goals, as well as the management of development opportunities and participation within the school.
Pedagogical Leadership: Involvement in the educational work	Confirmatory factor analysis	Frequency with which the principal monitors and supervises different aspects of teacher work (for example, planning activities, teaching practices, time use, and parent-teacher relationship).

Source: Ministry of Education and Vocational Training (2022b)

Analytical strategy

For data analysis, two stages were considered. The first focuses on performing a descriptive analysis of the characteristics of the selected sample. The second, given the nested structure of the data in an educational context (Gaviria & Castro, 2005; Raudenbush & Bryk, 2002), proposes three multilevel structure models to analyze the relationship that different factors have with ECE performance.

For this stage, a model that incorporated only the characteristics of the students was estimated. Then, based on the previous model, variables related to the teacher and her practice were included. Lastly, variables related to the principal were added. For the model adjustment, the AIC and BIC indicators were taken into account. These indicators were evaluated to decrease as the models became more complex. Analyses were performed using the *lmer* function of the R language package *lme4* (Bates et al., 2014).

The scales used were standardized with a mean of 0 and SD of 1. The multilevel models took the students and their characteristics as level 1 and the characteristics of the teachers and principals grouped in the classrooms as level 2. Being a representative sample of the urban area, weights built by the UMC were used for data analysis.

Results

The results obtained in this research are presented in two sections. Firstly, the distribution of the main characteristics of the sample is presented (Table V) and then the correlations between the variables used are analyzed. Second, models and the effects of variables on performance are examined. To do this, the fit indices of each model are compared.

Gender distribution is equal between men and women. Also, most of the students speak Spanish and most come from urban state schools. Regarding the socioeconomic status, 68.93% belong to the low and very low level. Only 7.06% belong to the high level. 6.51% of students have not attended preschool education (3 to 5 years) and 17.68% have repeated at least one grade.

Table VI shows the distribution of teachers and principals characteristics according to gender, experience, and educational level. It should be mentioned that 62.59% of the teachers are women. On the contrary, among principals they represent 32.57%. In addition, 37.39% of teachers have postgraduate studies. In the case of directors, 69.42% of them have a postgraduate degree. Finally, more than 80.00% of both groups are experienced professionals.

Figure I shows the correlations between the different student indices and performance in Reading. The diagonal presents the distribution of the index among the students. Below the diagonal, the data distribution

Table V. Distribution of students by sociodemographic features and educational trajectory

Features		Proportion (%)
Gender	Male	49.43
	Female	50.57
First language	Indigenous	8.59
	Spanish	91.41
Socioeconomic status	High	7.60
	Medium	23.47
	Low	33.04
	Very low	35.89
Preschool education	Attended	93.49
	Did not attend	6.51
Repetition	Repeated	17.68
	Did not repeat	82.32

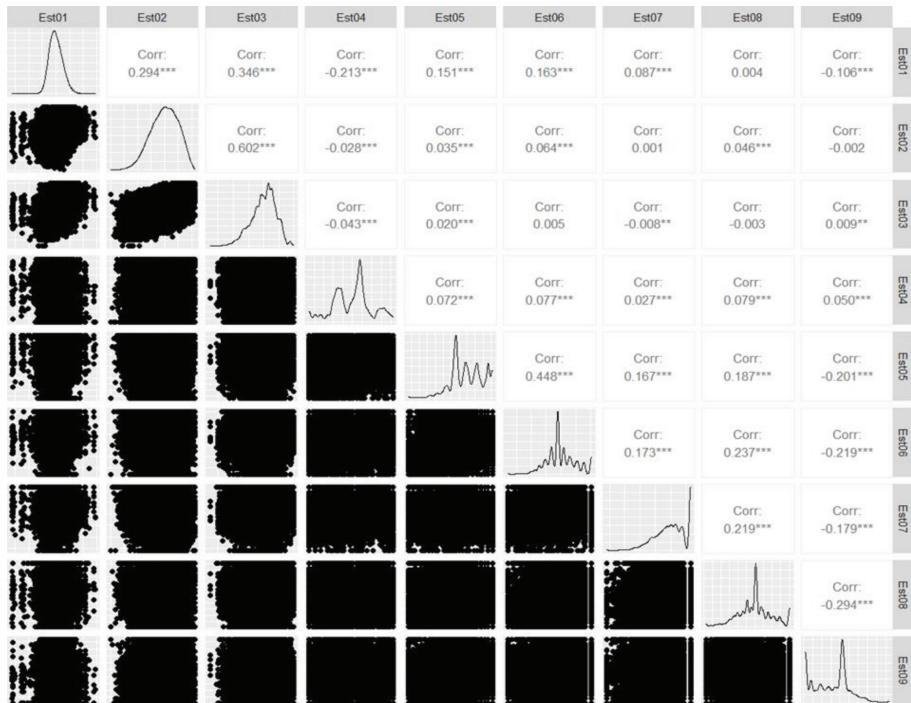
Source: Compiled by author

Table VI. Distribution of teachers and principals by features of gender and professional career

Feature		Teacher	Principal
Gender	Male	37.41	67.43
	Female	62.59	32.57
Teaching experience	Novice	14.67	19.33
	Experienced	85.33	80.67
Educational level Qualifications	College education at most	62.61	30.58
	Postgraduate (master's and/or doctorate)	37.39	69.42

Source: Compiled by author

Figure I. Correlations between Reading performance and student reported indices



***p<.001; **p<.01; *p<.05

Note: Est01: Reading Performance. Est02: Socioeconomic index of the student. Est03: Average socioeconomic index of students in the school. Est04: Beliefs about reading: reading as a natural talent. Est05: Beliefs about reading: reading as a learned competence. Est06: Reading self-efficacy. Est07: Pedagogical practices (general). Est08: Coexistence: relationship between students. Est09: Coexistence: sense of belonging.

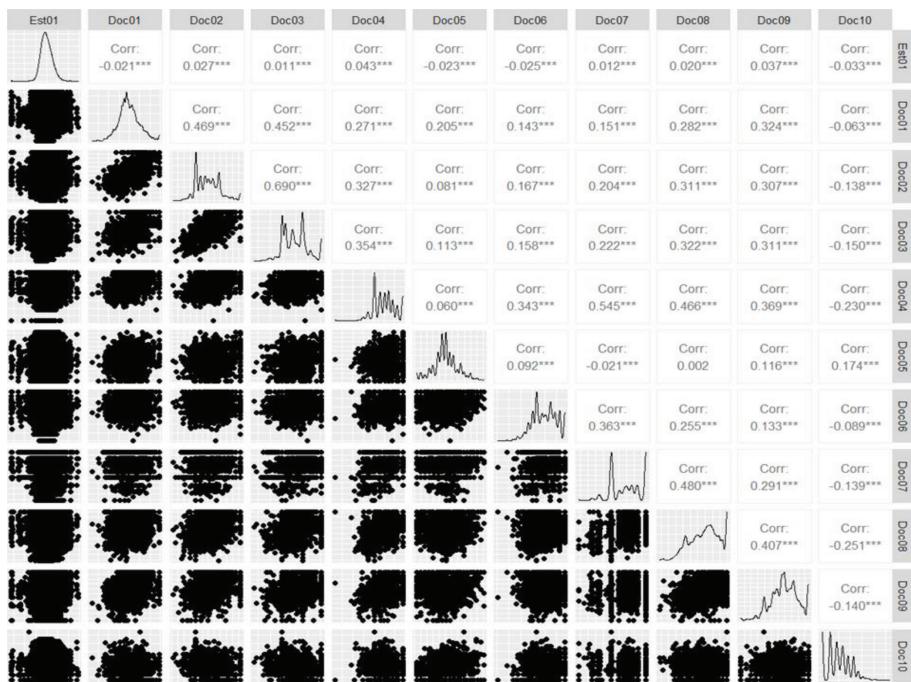
Source: Compiled by author

among the correlated variables is observed. Among the indices that most correlate with performance, the socioeconomic index (ISE) of the school (.346) stands out along with the individual student index (.294). “Coexistence: relationship between students” does not present a statistically significant correlation with performance and “beliefs about reading: reading as a natural talent” is negatively related to the latter.

The “coexistence: sense of belonging” index is negatively related to performance. This indicates that the higher the level of identification of the student with the school, the lower the performance tends to be. This relationship may be influenced by variables not considered in the study.

Regarding the relationship between the teacher's indices and performance, figure II shows that they all reach statistical significance. However,

Figure II. Correlations between Reading performance and teacher-reported index

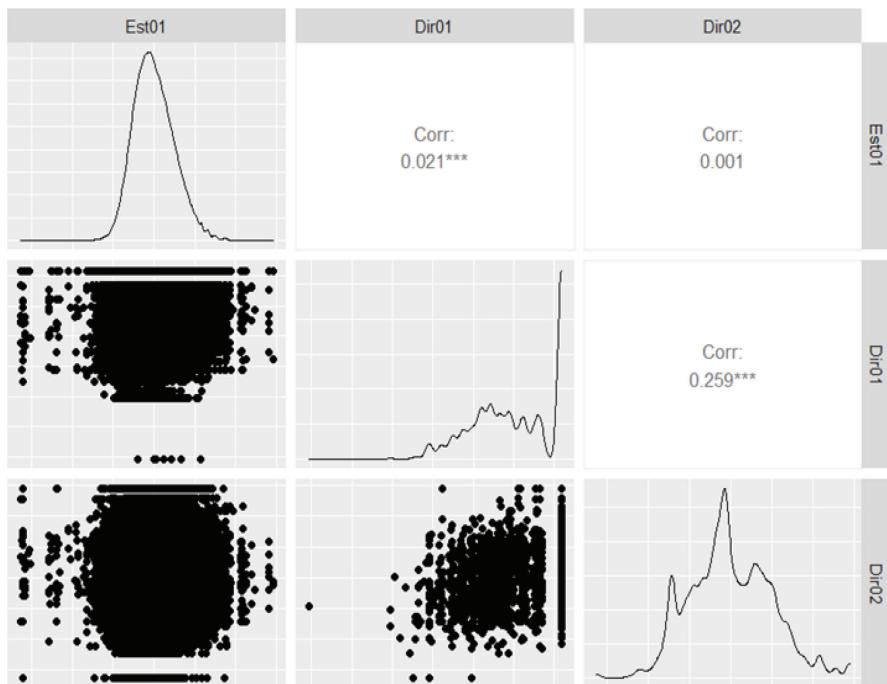


***p < 0,001; **p < 0,01; *p < 0,05

Note: Est01: Reading achievement. Doc01: Curricular coverage. Doc02: Reading different types of text. Doc03: Reading different text formats. Doc04: Self-efficacy on teaching reading. Doc05: Reading as a natural talent. Doc06: Reading as a learned competence. Doc07: Reading self-efficacy. Doc08: Reading pedagogical practices. Doc09: Interaction with other faculty. Doc10: Classroom discipline climate.

Source: Compiled by author

Figure III. Correlations between reading performance and principal-reported rates



***p < .001; **p < .01; *p < .05

Note: Est01: Reading Achievement. Dir01: Pedagogical leadership: strategic planning. Dir02: Pedagogical leadership: Involvement in the educational work.

Source: Compiled by author

the coefficients are quite small (less than .05). Likewise, three of the indices are negatively related to performance: curricular coverage, reading as a natural talent and disciplinary climate in the classroom. All reported indices present medium size correlations among them.

Figure III shows the indexes reported by the director. Despite reaching statistical significance, the strategic planning index presents a fairly small coefficient (less than .05).

Associated factor models

The results from the four models are presented in Tables VII and VIII. Table VII presents the effects at the level of student variables and Table

Table VII. Individual-level effects on reading performance

	Null model	Model 01	Model 02	Model 03
Interceptor	-0.07*** (0.01)	-0.13*** (0.01)	-0.11*** (0.02)	-0.16*** (0.02)
Student gender		0.03*** (0.01)	0.02*** (0.01)	0.02*** (0.01)
First Language		-0.35*** (0.01)	-0.26*** (0.01)	-0.26*** (0.01)
Socioeconomic index		0.12*** (0.00)	0.08*** (0.00)	0.08*** (0.00)
Preschool education		0.18*** (0.01)	0.17*** (0.01)	0.17*** (0.01)
Repitence		-0.33*** (0.01)	-0.33*** (0.01)	-0.33*** (0.01)
Reading as a natural talent		-0.18*** (0.00)	-0.18*** (0.00)	-0.18*** (0.00)
Reading as a learned competence		0.08*** (0.00)	0.08*** (0.00)	0.08*** (0.00)
Reading self-efficacy		0.11*** (0.00)	0.11*** (0.00)	0.11*** (0.00)
Perception on pedagogical practices		0.05*** (0.00)	0.05*** (0.00)	0.05*** (0.00)
Sense of belonging		-0.04*** (0.00)	-0.04*** (0.00)	-0.04*** (0.00)
Relationship between students		-0.07*** (0.00)	-0.07*** (0.00)	-0.07*** (0.00)
AIC	303821.71	287824.71	285922.55	285943.12
BIC	303850.64	287959.74	286192.61	286261.39
Log Likelihood	-151907.85	-143898.36	-142933.28	-142938.56
Num. obs.	114093	114093	114093	114093
Num. groups: classroom	5568	5568	5568	5568
Var. classroom (Intercepto)	0.24	0.12	0.08	0.08
Var. Residual	0.76	0.68	0.68	0.68

***p < .001; **p < .01; *p < .05

Source: Compiled by author

Table VIII. School-level effects on reading performance

	Null model	Model 01	Model 02	Model 03
Average socioeconomic index			0.23*** (0.01)	0.22*** (0.01)
Student gender			0.01 (0.01)	0.01 (0.01)
Teaching experience			-0.03* (0.01)	-0.03* (0.01)
Teachers's educational level			0.06*** (0.01)	0.06*** (0.01)
Curricular coverage.			-0.03*** (0.01)	-0.03*** (0.01)
Reading different types of text			0.03*** (0.01)	0.03*** (0.01)
Reading different text formats			-0.00 (0.01)	-0.00 (0.01)
Self-efficacy on teaching reading			0.02** (0.01)	0.02*** (0.01)
Reading as a natural talent			0.01 (0.00)	0.01 (0.00)
Reading as a learned competence			-0.00 (0.01)	-0.00 (0.01)
Reading self-efficacy			-0.02** (0.01)	-0.02*** (0.01)
Reading pedagogical practices			0.01 (0.01)	0.01 (0.01)
Interaction with other faculty			-0.02** (0.01)	-0.02** (0.01)

(Continued)

Table VIII. School-level effects on reading performance (Continued)

	Null model	Model 01	Model 02	Model 03
Classroom discipline climate			-0.02*** (0.00)	-0.02*** (0.00)
Principal gender				0.03*** (0.01)
Principal's educational level				0.02 (0.01)
Principal experience				0.04*** (0.01)
Pedagogical Leadership: Strategic Planning				0.00 (0.00)
Pedagogical Leadership: Involvement in the educational work				-0.01* (0.00)
AIC	303821.71	287824.71	285922.55	285943.12
BIC	303850.64	287959.74	286192.61	286261.39
Log Likelihood	-151907.85	-143898.36	-142933.28	-142938.56
Obs. num.	114093	114093	114093	114093
Gr. Num.: classroom	5568	5568	5568	5568
Var. classroom (Intercepto)	0.24	0.12	0.08	0.08
Var. Residual	0.76	0.68	0.68	0.68

***p < .001; **p < .01; *p < .05

VIII the effects at the level of the school context. In addition, the values of the AIC and BIC goodness-of-fit indicators for each model are shown. The first model proposed is the null model, with the objective of estimating how much of the variance of performance is explained by the characteristics of the student (level 1) and the school context (level 2) by calculating the intraclass correlation (ICC). The ICC obtained is .24, which means that the school context variables would explain 24.0% of the performance variance.

The first model proposed is the null model, with the objective of estimating how much of the variance of performance is explained by the characteristics of the student (level 1) and the school context (level 2) by calculating the intraclass correlation (ICC). The ICC obtained is .24, which means that the school context variables would explain 24.0% of the performance variance.

Model 01 considers variables related to the student. All variables are statistically significant. The sex of the student does not present a significant effect ($\beta_i=.03$). First language and repetition are relevant factors in their relationship with performance. Those students with a first original language obtain 35.0% of one SD less performance than their peers with Spanish as their mother tongue. Students who have repeated a grade obtain 33.0% of one SD less performance than those who have not repeated. Attending preschool education is related to a change in performance of 18.0% of a SD. However, the student's ISE presents a small coefficient ($\beta_i=.12$), possibly due to the socioeconomic homogeneity of the sample used (table VIII).

The indices of perception of ability in reading and self-efficacy for reading have a negative effect on performance, with changes of 18.0% and 11.0% of one SD, respectively. However, it should be noted that performance also affects these scales. Those students with high achievement could present high levels of self-efficacy in reading. For these indices the correlation between both variables is small (less than .20).

The perception of students about the teacher's pedagogical practices has a low influence on performance ($\beta_i=.05$). Perhaps due to desirability problems in the students' response. An indication of this is the accumulation of cases to the right in the scale density (figure I), which implies little variability of the scale. The coefficients of these variables are similar in the following models, except for the student's ISE, which is reduced by including the school's average ISE.

The following models included school context variables, such as the teacher's pedagogical practice (model 02) and, later, principal variables (model 03). The results (table VIII) show that the school's average ISE has a medium impact on performance ($\beta_i=.23$). This means that the schools with the highest average ISE achieve a higher performance in Reading. This could be related to conditions of socioeconomic segregation within the educational system. Variables such as teacher experience, educational level, curricular coverage, reading of various types of text, self-efficacy about teaching reading, self-efficacy about reading, interaction

with other teachers, and the climate of discipline in the classroom turn out to be statistically significant, but with coefficients too small to have any effect on performance. In model 03, none of the direct variables and their school management were relevant to student performance.

In relation to the goodness of fit of these models, it is observed that model 02 improves the values of the AIC and BIC compared to model 01, but the addition of variables related to the director and his management (model 03) do not contribute to the performance explanation.

Related to the fit goodness of these models, it is observed that model 02 improves the values of the AIC and BIC compared to model 01, but the addition of variables related to the director and his management (model 03) do not contribute to the performance explanation.

The results indicate that the socio-familiar context is the main factor influencing performance in Reading, especially grade repetition and the average ISE of the school. Teacher-related factors, such as their pedagogical practice and school management, are not related to reading performance. These results are contradictory to what was expected and raise the question of why teachers' pedagogical practices are not related to student achievement. Possible explanations will be discussed in the next section.

Discussion

The research sought to establish a relationship between the teacher's pedagogical practice and the development of students' competencies in the area of Communication, through the use of multilevel regression models with data from the ECE of 2019. Taking this into account, below, is presented the discussion of the findings, the limitations of the study and some lines of research that in the future could complement what is stated in this document.

The results of the ECE 2019 in Reading among urban public school students are mainly associated with individual and attitudinal factors

The results indicate that student performance is mainly linked to the socioeconomic and family context, with students in a situation of economic vulnerability having fewer learning opportunities. This indicates

that urban public schools are not succeeding in overcoming the socio-economic inequalities of their contexts. Thus, the evidence suggests that performance is determined by the socioeconomic starting point of the student, raising questions about how the school can deal with this inequality and what aspects need to be improved in teachers to achieve significant improvement.

Likewise, it is striking that the student's sex is not associated with differences in performance. Several investigations indicate that being a woman is usually associated with higher performance in the Communication area. On the contrary, this is not the case among Peruvian urban public school students. However, this could change if students from private and/or rural schools were considered in the analysis. In any case, these schools seem to be developing a similar amount of learning between male and female students. In line with the above, the size of the repetition effect is also striking. This invites us to reflect on the need to analyze in depth the relationship between both variables and the sense of causality that could exist between them (López & García, 2021).

Within the variables related to the perceptions of the students, the findings point to the importance of the belief about the nature of reading ability as something innate. The relationship of this variable with performance can be explained by a greater exposure to learning opportunities that students would be having, both inside and outside of school. Students exposed to more reading tasks are likely to perceive improvements in their reading ability and academic performance in this area. This shapes and feeds back the perception that reading skills are learned and modified.

The null relationship between the pedagogical practices of the teacher and the pedagogical leadership of the director could be related to factors of the Peruvian educational context and problems of desirability in the information report

The fact that the variables related to the teacher and their pedagogical practice do not have an impact on performance makes sense given the overload of information that teachers receive from different fronts. On the one hand, there are the regulations, supporting documents and training that the Ministry of Education has developed and makes available to teachers and, on the other hand, the requirements and training of the local educational authorities

(regional education directorates and the local educational management units). Additionally, teachers must deal with the particular demands that educational institutions place on them. This leaves little time for teachers to process and apply this information in the classroom.

Also, the educational reforms promoted by the State to improve teaching quality in schools have failed due to the lack of articulation between the different levels of educational management and the constant changes in the approaches to interventions (Cuenca & Vargas, 2018). Likewise, teachers have not been able to appropriate curricular documents and effective pedagogical practices (Guerrero, 2009; León Zamora, 2018; Rivas, 2015). On the other hand, it should be considered that the collection of perceptions through multiple-choice questionnaires has a limitation associated with the difficulty of collecting honest responses due to desirability.

This result calls for reflection on the distance between what the teacher reports about his pedagogical practice and what happens in the classroom. Also, how much is known about the process that the teacher follows to materialize his pedagogical practice in class. In addition, it must be considered that these students have gone through school during a period of constant curricular changes (Cuenca, 2013; León Zamora, 2018; Rivas, 2015) and their learning may have been affected. In this sense, political instability and the lack of articulation between the different levels of the State can be a negative factor in the educational trajectory of students (de Belaunde, 2011; Matos Mar, 2011).

The results of the multilevel models show a low relationship between the variables associated with the teacher and academic performance. This indicates that the effectiveness of the teacher is not framed in a single aspect but in several related ones. Therefore, any action or intervention to improve teacher quality must consider the perceptions and beliefs of teachers and not only their knowledge or certification, both in their initial training and in service. It is also important to take into account that the questionnaire used considers certain pedagogical practices and it is possible that others not considered also have an effect on learning.

Study limitations and future lines of research

At this point, it is important to mention the limitations of this investigation. Despite having proposed models that cover several aspects of the students' socio-educational context, this area remains extremely complex.

For this reason, further efforts should be made to identify ways to obtain objective and reliable measures that reflect the actual pedagogical practices of teachers.

The findings of this research suggest the need to carry out qualitative studies on the mechanisms used by teachers in their daily pedagogical practices in the classroom. In addition, having a system of objective observation in the classroom and analyzing contextual factors of the home such as the involvement of parents and the resources for learning to read within the home.

Bibliographic references

Arcia, G., & Laguna, J. (2004). *Análisis de los factores asociados con el rendimiento académico en 3º y 6º grados de primaria*. Ministerio de Educación de Nicaragua.

Arenas, C., & Ámbros, A. (2021). Efecto mediacional de las habilidades lectoras básicas e intermedias en la relación comprensión y rendimiento académico de estudiantes de primero de Educación Secundaria Obligatoria. *Revista de Educación*, 394, 129-156. DOI: 10.4438/1988-592X-RE-2021-394-503

Ausbel, D. (1976). *Psicología educativa. Un punto de vista cognoscitivo*. Editorial Trillas.

Balarin, M. (2016). La privatización por defecto y el surgimiento de las escuelas privadas de bajo costo en el Perú. ¿Cuáles son sus consecuencias? *Revista de Sociología de la Educación-RASE*, 9(2), 181-196.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2014). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. DOI: 10.18637/jss.v067.i01

Belfi, B., Gielen, S., Fraine, B., Verschueren, K., & Meredith, C. (2015). School-based social capital: The missing link between schools' socio-economic composition and collective teacher efficacy. *Teaching and Teacher Education*, 45, 33-44. DOI: 10.1016/j.tate.2014.09.001

Beltrán, A., & Seinfeld, J. (2012). *La trampa educativa en el Perú; Cuando la educación llega a muchos, pero sirve a pocos*. Universidad del Pacífico.

Berger, M., & Toma, E. (1994). Variation in State Education Policies and Effects on Student Performance. *Journal of Policy and Management*, 13(3), 477-491.

Campos, D., Contreras, P., Riffó, B., Véliz, M., & Reyes, A. (2014). Complejidad textual, lecturabilidad y rendimiento lector en una prueba de comprensión en escolares adolescentes. *Universitas Psychologica*, 13(3), 1135-1146.

Caro, D. (2003). ¿Cómo mejorar el desempeño académico de los estudiantes que asisten a escuelas en las zonas pobres del Perú? En E. Vásquez y D. Winkelried (Eds.), *Buscando el bienestar de los pobres: ¿Cuán lejos estamos?* (pp. 175-207). Universidad del Pacífico.

Caro, D., Lenkeit, J., & Kyriakides, L. (2015). Instructional approaches and differential effectiveness across learning contexts: Evidence from PISA 2012. *Studies in Educational Evaluation*, 49, 30-41.

Coleman, J. (1966). *Equality of Educational Opportunity*. U.S. Department of Health, Education and Welfare. Office of education.

Creemers, B., & Kyriakides, L. (2007). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge. DOI: 10.4324/9780203939185

Cuenca, R. (2013). La escuela pública en Lima Metropolitana ¿Una institución en extinción? *Revista Peruana de Investigación Educativa*, 5, 73-98.

Cuenca, R., & Vargas, J. (2018). *Perú: el estado de las políticas públicas docentes (Documento de trabajo, 242. Serie Educación, 17)*. Instituto de Estudios Peruanos.

Cueto, S., León, J., Ramírez, C., & Guerrero, G. (2008). Oportunidades de aprendizaje y rendimiento escolar en Matemática y Lenguaje. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 6(1), 29-41.

De Belaunde, C. (2011). Profundizando las brechas. Una mirada a la desigualdad en los estudios sobre el sistema educativo peruano. En J. Cotler y R. Cuenca (Eds.), *Las desigualdades en el Perú. Balances críticos* (pp. 273-329). Instituto de Estudios Peruanos

Ertmer, P., & Newby, T. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2), 43-71

Freire, S., & Miranda, A. (2014). *El rol del director en la escuela: el liderazgo pedagógico y su incidencia sobre el rendimiento académico (Avances de investigación 17)*. Grupo de Análisis para el Desarrollo.

Fuller, B., & Clarke, P. (1990). Raising schools effects while ignoring culture. Local conditions and the influence of classroom, tools, rules and pedagogy. *Review of Educational Research*, 64(1), 119-157.

Gaviria, J., & Castro, M. (2005). *Modelos jerárquicos lineales*. La Muralla.

Guerrero, L. (2009). Política docente: balance del período. En Foro Educativo (Ed.), *La educación en los tiempos del APRA: balance 2006 - 2009* (pp. 15-36). Foro Educativo.

Hammonds, C., & Lamar, C. (1972). *La enseñanza: su orientación, sus funciones, sus motivaciones*. Editorial Trillas.

Hanushek, E., & Taylor, L. (1990). Alternative Assessments of the Performance of Schools: Measurement of State Variations in Achievement. *Journal of Human Resources*, 25(2), 179-201.

Hanushek, E., & Woessmann, L. (2009). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267-321. DOI: 10.1007/s10887-012-9081-x

Harbinson, R., & Hanushek, E. (1992). *Educational performance of the poor: lessons from rural northeast Brazil*. World Bank Group.

Hernández, P., Hernández, C., Capote, C., & García, J. (2003). *La percepción del alumnado de las teorías psicoeducativas de los mejores y peores profesores* [Comunicación presentada en el IV Congreso Internacional de Psicología y Educación]. Recuperado de <https://www.tafor.net/images/Comunica001.pdf>.

Horn, A., & Marfán, J. (2010). Relación entre liderazgo educativo y desempeño escolar: Revisión de la investigación en chile. *Psicoperspectivas*, 9(2), 82-104. DOI: 10.5027/PSICOPERSPECTIVAS-VOL9-ISSUE2-FULLTEXT-116

La Serna, K. (2011). El producto del aprendizaje: el rendimiento académico. En K. La Serna (Ed.), *Retos para el aprendizaje: de la educación inicial a la universidad. Investigaciones explicativas sobre el rendimiento académico de los estudiantes peruanos* (pp. 19-103). Universidad del Pacífico.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management*, 28(1), 27-42.

León Zamora, E. (2018). *El fenómeno ECE y sus efectos en las prácticas docentes*. ENACCIÓN.

Lloyd, P., & Fernyhough, C. (1999). *Lev Vigotsky. Critical Assessments* [Vol. I]. Routledge.

López, F., & García, I. (2021). La repetición de curso y la graduación en Educación secundaria obligatoria en España. Análisis empíricos y recomendaciones políticas. *Revista de Educación*, 394, 325-353. DOI: 10.4438/1988-592X-RE-2021-394-510

Martín, E., Marchesi, A., & Pérez, E. (2008). Variables that Predict Academic Achievement in the Spanish Compulsory Secondary Educational System: A Longitudinal, Multi-Level Analysis. *The Spanish Journal of Psychology*, 11(2), 400-413. DOI: 10.1017/S113874160000442X

Martín, L., Aramayo, K., Rodríguez, J., & Saiz, M. (2022). La procrastinación en la formación inicial del profesorado: el rol de las estrategias de aprendizaje y el rendimiento académico. *Educación XXI*, 25(2), 65-88. DOI: 10.14201/teri.23580

Martínez, T., Ávila, V., Ysla, L., & Sellés, P. (2022). Fortalecimiento de los precursores de la Lectura en estudiantes prelectores. *Revista de Educación*, 396, 177-204. DOI: 10.4438/1988-592X-RE-2022-396-534

Matos Mar, J. (2011). *Estado desbordado y sociedad nacional emergente*. Universidad Ricardo Palma.

Murillo, J., & Martínez, C. (2018). Factores de aula asociados al desarrollo integral de los estudiantes: un estudio observacional. *Estudios pedagógicos XLIV*, 1, 181-205.

Ministry of Education (2016a). *National curriculum for basic education*. Author.

Ministry of Education (2016b). *Regions in perspective. The influence of factors associated with learning at the end of primary education* [Estudios Breves N.º 2]. Office of Learning Quality Measurement.

Ministry of Education. (2017) *What happens to learning in the transition from primary to secondary school? A look from equity and differentiated school efficiency*. Office of Learning Quality Measurement.

Ministry of Education (2018). *Profiles of pedagogical leadership in primary schools and their relationship with performance* [Zoom educativo N.º 5]. Office of Learning Quality Measurement.

Ministry of Education. (2020a). *ECE - EM 2019 results presentation* [PowerPoint presentation]. Retrieved from <http://umc.minedu.gob.pe/resultadosnacionales2019/>.

Ministry of Education. (2020b). *Technical report on student, teacher, and principal indices in the 2019 Student Census Assessment* [Internal working document]. Office of Learning Quality Measurement.

Ministry of Education. (2022a). *Understanding the focus of the Communication area and the CNEB competencies* [PowerPoint presentation]. Retrieved from <https://minedu.digital/?s=enfoque+del+area+de+comunicación>.

Ministry of Education. (2022b). *Matrix of questionnaire indices for the 2019 Student Census Assessment [Internal working document]*. Office of Learning Quality Measurement.

Mizala, A., & Romaguera, P. (1999). *Factores que determinan el desempeño escolar en Bolivia [Documento de trabajo N.º 61]*. Centro de Economía Aplicada de la Universidad de Chile.

Monge, C., & Gómez, P. (2020). El papel de la convivencia escolar en la formación inicial del profesorado de educación infantil y primaria. *Teoría de la Educación. Revista Interuniversitaria*, 33(1), 197-220. DOI: 10.14201/teri.23580

Ortega, P., & Romero, E. (2020). El valor de la experiencia del alumno como contenido educativo. *Teoría de la Educación. Revista Interuniversitaria*, 33(1), 89-110. DOI: 10.14201/teri.23615

Phillips, D. (1995). The good, the bad and the ugly: the many faces of constructivism. *Educational Researcher*, 24(7), 5-12.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. SAGE Publications. DOI: 10.2466/pms.2002.94.2.671

Risso, A., Peralbo, M., & Lozano, A. (2010). Cambios en las variables predictoras del rendimiento escolar en enseñanza secundaria. *Psicothema*, 22(4), 790-796.

Rivas, A. (2015). *América Latina después de PISA: lecciones aprendidas de la educación en siete países (2000-2015)*. CIPPEC-Natura-Instituto Natura.

Rodríguez, M., Jenaro, C., & Castaño, R. (2022). La percepción de los alumnos como indicador de inclusión educativa. *Educación XXI*, 25(1), 357-379. DOI: 10.5944/educXXI.30198

Serván, S., & Tantaleán, E. (2011). ¿Explican la formación y las estrategias didácticas del docente las diferencias del rendimiento académico de los estudiantes de colegios públicos y privados? En K. La Serna (Ed.), *Retos para el aprendizaje: de la educación inicial a la universidad. Investigaciones explicativas sobre el rendimiento académico de los estudiantes peruanos* (pp. 171-249). Universidad del Pacífico.

Summer, A., & Wolfe, B. (1977). Do Schools Make a Difference? *American Economic Association*, 67(4), 639-652. Recuperado de <https://www.jstor.org/stable/1813396>.

Von Glaserfeld, E. (1996). Footnote to “The many faces of constructivism”. *Educational Researcher*, 25(6), 19.

Willms, D., & Somer, M. (2001). Family, Classroom, and School Effects on Childrens Educational Outcomes in Latin America. *School Effectiveness and School Improvement*, 12(4), 409-445. DOI: 10.1076/sesi.12.4.409.3445

Contact address: Manuel Alfredo Marcos Balabarcá. Universidad Complutense de Madrid, Facultad de Educación y Formación del Profesorado c/ Rector Royo Villanova, s/n . 28040, Madrid, Spain. E-mail: manumarc@ucm.es

A Bibliometric Review of Instructional Leadership Research: Science Mapping the Literature from 1974 to 2020

Una revisión bibliométrica de la investigación sobre liderazgo educativo: Mapeo científico de la literatura de 1974 a 2020

<https://doi.org/10.4438/1988-592X-RE-2023-401-592>

Chuan-Chung Hsieh

<https://orcid.org/0000-0002-9815-7350>

National Tsing Hua University

Imam Gunawan

<https://orcid.org/0000-0003-3462-0729>

National Tsing Hua University

Universitas Negeri Malang

Hui-Chieh Li

<https://orcid.org/0000-0002-4474-8235>

National Taipei University of Business

Abstract

This study used bibliometric science mapping to explore the research development status and intellectual structure of instructional leadership and to identify research fronts and hotspots in instructional leadership studies. Relevant citation data screened from Web of Science revealed 1172 records spanning from 1974 to 2020, which were analyzed using HistCite™, VOSviewer and Sci2 Tool. Descriptive statistics revealed four development stages along with significant articles in each stage. Document bibliographic coupling and content analyses indicated eight major research clusters with their respective research focus. Burst detection and keywords co-occurrence analyses identified research fronts including shared leadership, teaching strategies, systematic review, principal preparation,

and school climate. Comparison of analysis results obtained using different tools showed discrepancies, thus highlighting the need for different analytical tools to be adopted as they complement each other in offering multiple and complementary perspectives for an across-the-board overview. Finally, implications and limitations of this study are presented.

Keywords: instructional leadership, schools' leadership, principal, science mapping, bibliometric analysis

Resumen

Este estudio utilizó el mapeo científico bibliométrico para explorar el estado de desarrollo de la investigación y la estructura intelectual del liderazgo educativo y para identificar frentes de investigación y puntos críticos en los estudios de liderazgo educativo. Los datos de citas relevantes seleccionados de Web of Science revelaron 1172 registros que abarcan desde 1974 hasta 2020, que se analizaron con HistCite™, VOSviewer y Sci2 Tool. Las estadísticas descriptivas revelaron cuatro etapas de desarrollo junto con artículos significativos en cada etapa. El acoplamiento bibliográfico de documentos y los análisis de contenido indicaron ocho grupos principales de investigación con sus respectivos focos de investigación. Los análisis de detección de ráfagas y co-ocurrencia de palabras clave identificaron frentes de investigación que incluyen liderazgo compartido, estrategias de enseñanza, revisión sistemática, preparación del director y clima escolar. La comparación de los resultados de los análisis obtenidos con diferentes herramientas mostró discrepancias, lo que destaca la necesidad de adoptar diferentes herramientas analíticas, ya que se complementan entre sí al ofrecer perspectivas múltiples y complementarias para una visión global. Finalmente, se presentan las implicaciones y limitaciones de este estudio.

Palabras clave: liderazgo educativo, liderazgo escolar, director, mapeo científico, análisis bibliométrico

Introduction

Instructional leadership (IL) has been defined as actions directly related to teaching and learning that aim to improve teaching tools and methods by initiating reflection and influencing teacher goals, values, and practices (Leithwood & Duke, 2009). Both teacher- and student-centered, IL targets at delivery of quality instruction (Juma et al., 2021). As instructional leaders, principals influence classroom teaching through formulating

school goals, setting and communicating achievement expectations, organizing classrooms, allocating resources, assessing teacher performance, evaluating student learning progress, and creating a positive and orderly school environment for learning (Heck et al., 1990). In essence, IL is the leadership behavior of the principal in influencing the learning process; therefore, its focus is on the actions taken by the principal to improve instructional quality.

Empirical research on IL includes the pioneering work of Edmonds (1979) and recent studies of Skaalvik (2020), which investigated the significant influence of IL on teacher professional development, motivation, and job satisfaction. Reitzug et al. (2008) proposed four dominant conceptions, namely relational, linear, organic, and prophetic IL, and discussed their implications for research and practice. Cale et al. (2015) critically explored IL in the context of special education in small to medium town schools. They identified a set of factors including communication, teacher evaluation and supervision, staff development, instructional programming, and instructional design that were crucial to the implementation of IL. Day et al. (2016) examined both direct and indirect impacts of principals applying both transformational and IL on student outcomes.

On the basis of content analysis results, Rigby (2013) proposed three logics of IL, namely prevailing logic, entrepreneurial logic, and social justice logic. The IL framework proposed by Hallinger and Murphy (1985) comprises three dimensions: defining the mission of the school, managing instructional programs, and promoting the school learning climate. In addition to these, Weber (1996) identified two more dimensions of IL, which include observing and enhancing teaching quality, and evaluating programmed teaching.

Focusing on school leadership relations between principals and teachers, Marks and Printy (2003) evaluated the potential of their active collaboration around instructional matters and found substantial effect IL on school performance, measured by the quality of its pedagogy and the achievement of its students. Wahlstrom and Louis (2008) further explored the role of the professional community of teachers, which aimed at reducing teachers' dependence on principals as instructional leaders. Their study found that only when the professional community was weak did teachers turn to principals for direct instructional support. In addition to academic research, there are books that specifically investigate and promote IL (Hallinger et al., 2015; Townsend, 2019; Weber, 1996).

This study conducted a systematic literature review on IL with data obtained from the Web of Science (WoS) database. To identify leading research in the field of IL, descriptive analysis, document bibliographic coupling analysis, content analysis, keyword co-occurrence analysis and burst detection analysis were performed. Thus, the analysis results obtained would present a perspective different from conventional literature reviews, because the study with a systematic literature review are quite comprehensive as well as less biased and more transparent that allow large data sets to be represented meaningfully (Meza, 2021). Moreover, the knowledge constructs of IL are linked and visualized with network analysis in the form of clusters and networks. Furthermore, related literature included in the analysis covered almost half a century, from 1974-2020, which would shed light on the evolution of IL research over time. The research questions examined are as follows:

- What is the volume and growth trajectory of the IL journal literature?
- Which authors and documents have the greatest influence on IL?
- What are the most popular topics investigated by the IL scholars?
- How have research themes evolved over time and what are the current research fronts?

Method

This systematic review of research used bibliometric analysis to gain insights into the key documents and topics on IL research. Bibliometric complements traditional reviews and meta-analyses that look objectively at a particular area of control sign, for example a specific time frame or a limited sample of journals, to assess the productivity and frequency of scientific work, and word frequency (Pritchard, 1969). In recent years, bibliometric analysis has been a popular method increasingly used in the scientific community. Results of bibliometric analyses in this research can be of use to scholars in understanding current status and identifying future research opportunities in the field of IL. Methods used in this study for exploring the knowledge domains of IL research include descriptive analysis, document bibliographic coupling analysis, content analysis, keyword co-occurrence analysis and burst detection analysis.

Bibliographic coupling occurs when two documents both cite one or more documents likewise. The more citations to other documents they share, the higher their coupling strength. Capable of identifying 'hot' research topics, bibliographic coupling relies on appropriate thresholds set for number of related documents and the strength of bibliographic links (Glänzel & Czerwon, 1996). Content analysis aims for the subjective interpretation of the text data through the systematic classification process of coding and identifying themes or patterns (Hsieh & Shannon, 2005).

Keyword co-occurrence analysis explores links between keywords to understand the knowledge components and knowledge mapping of a scientific field (Radhakrishnan et al., 2017). Visual representation of co-occurrence networks shows nodes of keywords representing the cumulative knowledge of a domain, and links denoting co-occurrence of word pairs. Link weights are calculated according to the number of times a pair of words appear together in documents. Burst detection analysis, proposed by Kleinberg (2003), identifies time periods in which a target event is uncharacteristically frequent, or "bursty". To identify the research fronts of IL, this study analyzed the average year of publication for keywords, supplemented by keyword burst detection analysis, using Kleinberg's algorithm, to identify topics showing significant change of research interest. Such analysis brings to light both topics that have received attention over a short period but then lost favor, as well as current research fronts in the burst period including the present.

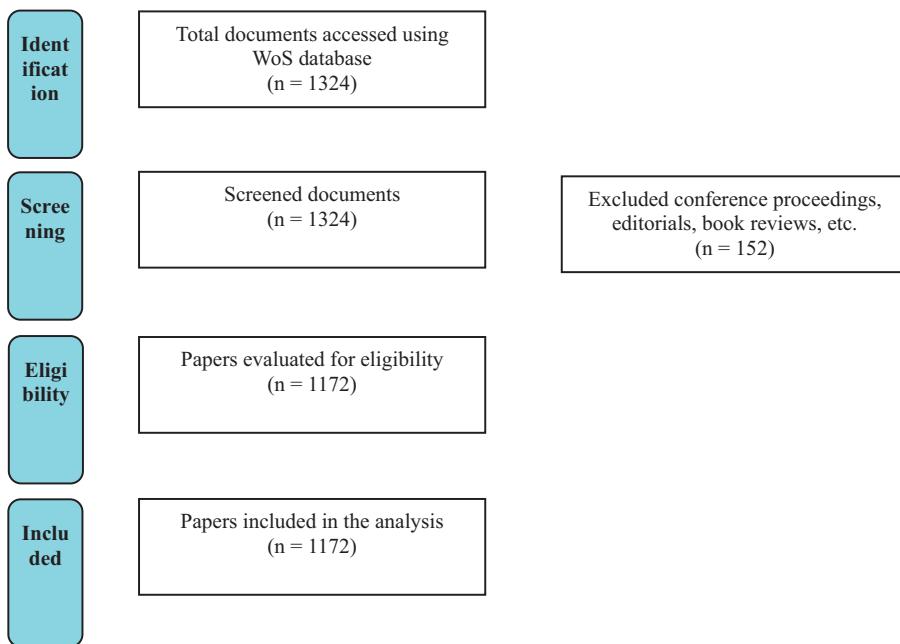
Data source, procedure, and analytic software

Data analyzed in this study were extracted from the Social Sciences Citation Index (SSCI), Science Citation Index Expanded (SCI-Expanded), and Arts & Humanities Citation Index (A&HCI) databases in the WoS Core Collection, which is a common source for bibliometric research. WoS includes the most reliable, high-impact scientific studies (Zyoud et al., 2017), and leading scientific citation search and analytical information platform supporting diverse scientific tasks across multiple knowledge domains as well as a dataset for large-scale data-intensive studies (Li et al., 2018). Moreover, about 99.11% of the journals indexed in the WoS database are also indexed in the Scopus database (Singh et al., 2021). This

study also used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Figure I) which provide a transparent and standardized scheme to visualize the identification and selection of study results in the bibliometric review and meta-analysis process (Moher et al., 2009). Related literature was identified from the WoS database using “instructional leadership” for search in terms of “Topic”. The categories chosen were Education and Educational Research, with a time span of 1974 to 2020, and a total of 1324 studies were screened. After excluding conference proceedings, editorial materials, and book reviews and chapters, the search performed in April 2021 yielded a total of 1172 documents (Figure I).

Data collected were processed for knowledge mapping using three bibliometric analysis and information visualization tools, namely HistCite™, VOSviewer and Sci2 Tool. The HistCite™ software analyzed inputs

FIGURE I. Flow diagram of study selection process



Source: Compiled by author

in the form of bibliometric records on co-citations of scientific articles (Barreiro, 2015). The VOSviewer software analyzes complex networks with its own group analysis function according to the strength of the connection between one project and another (van Eck & Waltman, 2020). On one hand, VOSviewer processed the data collected on the basis of co-occurrence; on the other hand, it generated network maps for result visualization. Sci2 Tool is a modular toolset specifically designed for the study of science (Sci2 Team, 2009), and can load data sets in different formats to conduct fundamental analysis such as burst detection analysis, co-occurrence, and coupling analysis.

Results and discussion

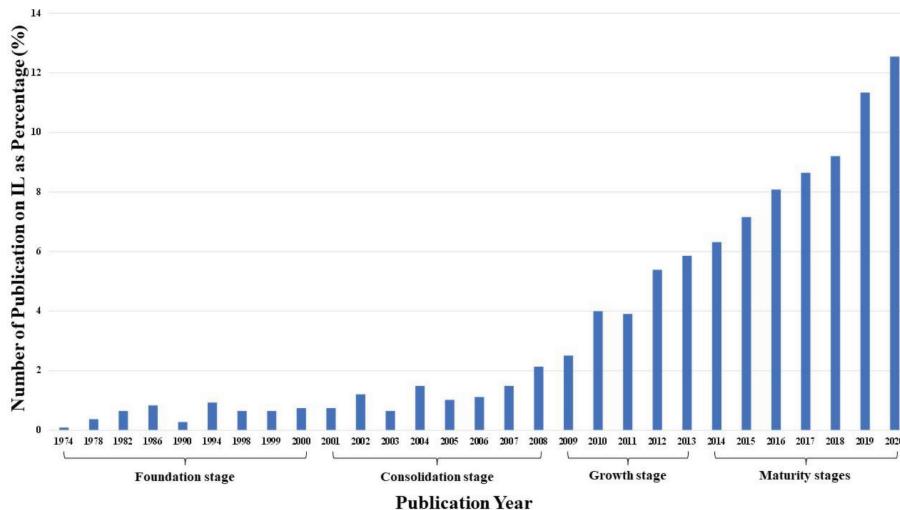
Yearly quantitative distribution of literature

Results of descriptive analysis shown in Figure II illustrate changes in the number of publications on IL between 1974 and 2020. Over this almost half a century, four stages of development can be identified. From 1974 to 2000, it was the Foundation stage, during which there were few publications and the number of publications each year showed no significant difference, on average six per year. Then came the Consolidation stage between 2001 and 2008 with the number almost doubled, albeit less than 20 per year. The average number of publications increased significantly to 50 per year in the Growth stage that followed, spanning from 2009 to 2013. From 2014 till present was the Maturity stage with an ever-increasing number of publications per year, two-fold that of the Growth stage. Figure II illustrates the number of publications on IL as a percentage. As can be seen, the number of publications in both Foundation and Consolidation stages are relatively few and insignificant, but increase sharply in both Growth and Maturity stages.

Significant publications in different development stages identified by Histcitetm

As Table I shows, Local Citation Score (LCS) is indicative of the citation frequency of a publication in the collection. The higher the LCS, the more

FIGURE II. Number of publications on IL as percentage between 1974 and 2020



Source: Compiled by author

frequent it is cited and the more significant it is in the research domain. Take Robinson et al. (2008) for example, it has the highest LCS of 131 among those listed in Table I, meaning that it is most cited in research publications on IL. Along with Robinson et al. (2008), publications of the Consolidation stage have much higher LCS than those in other stages, indicating their significant influence in promoting further development in the field of IL. Papers with comparatively low LCS, including Grissom et al. (2013) and those of the Maturity stage, are more recent publications of the past decade and it would take time for them to accumulate citations, and their impact on subsequent development of IL research is yet to be seen.

As Table I shows, during the Foundation stage, definitions and concepts of IL were still vague. The main papers published during this period (Blase & Blase, 1999; Hallinger & Murphy, 1985; Heck et al., 1990) focused on investigating, exploring, and defining the concept of IL, formulating dimensions and behaviors of IL, and exploring instructional management.

TABLE I. Significant publications on IL in different development stages identified by HistCite™

Period	Rank	Author(s)/Year	LCS
Foundation stage (1974-2000)	1	Hallinger and Murphy (1985)	93
	2	Blase and Blase (1999)	54
	3	Heck et al. (1990)	44
Consolidation stage (2001-2008)	1	Robinson et al. (2008)	131
	2	Marks and Printy (2003)	116
	3	Spillane et al. (2004)	40
Growth stage (2009-2013)	1	Supovitz et al. (2010)	66
	2	Neumerski (2012)	46
	3	Grissom et al. (2013)	26
Maturity stage (2014-present)	1	Shatzer et al. (2014)	24
	2	Day et al. (2016)	24
	3	Goddard et al. (2015)	21

Note: The article marked in gray also appears in Table III

Source: Compiled by author

The chief emphasis in this period was on refining the principles and concepts of IL. Comparison was made between IL and other leadership models in terms of effectiveness in improving school outcomes. The meta-analysis of Robinson et al. (2008) found IL three to four times more effective than transformational leadership in enhancing student academic and non-academic outcomes. Their findings were consistent with those reported by Marks and Printy (2003) that IL contributed more to improving school performance than transformational leadership.

Significant development during the Growth stage saw empirical studies conducted on assessing the contribution of IL to teaching performance and learning achievement (Supovitz et al., 2010). In addition, Neumerski (2012) further reviewed IL of principals, teachers and coaches as well as their interaction with followers when they work toward the improvement of teaching and learning. Complementing comparison of different leader types, the study of Grissom et al. (2013) with a unique data source of in-person, full-day observations collected over three years offered longitudinal evidence on the effective use of instructional time of school principals.

Moreover, in the early 2000s during both Consolidation and Growth stages, scholars began to re-conceptualize IL more broadly, as evidenced by the emergence of “shared instructional leadership” (Marks & Printy, 2003), “teacher leadership” (York-Barr & Duke, 2004), and “leadership for learning” (Murphy et al., 2007). These leadership models reframed IL as a distributed process that not only focuses on student learning, but also enhances teacher capacity and teacher commitment as well as designs school organizations to achieve their main goals.

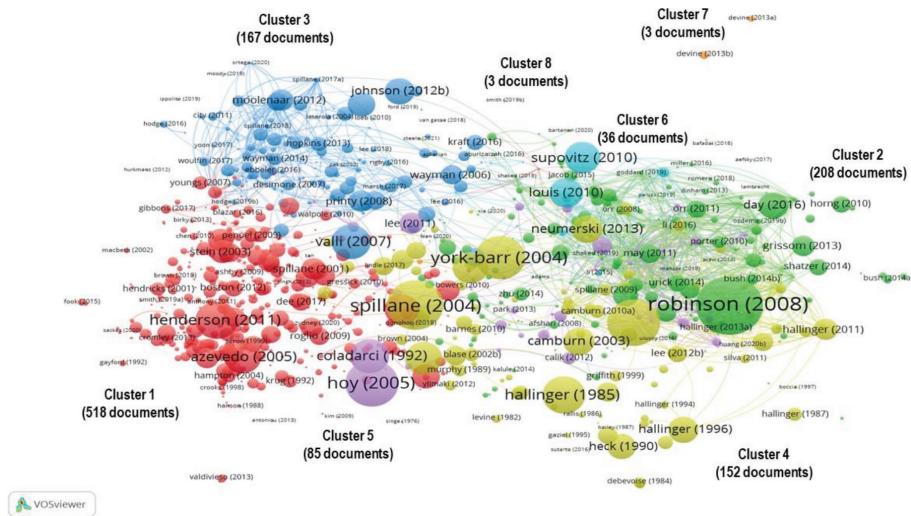
From the Maturity stage till the present, the research focus of IL has shifted towards massive integration and comparative study on impact of instructional and transformational leadership on student achievement (Day et al., 2016; Shatzer et al., 2014). Aiming for a more comprehensive exploration, Goddard et al. (2015) conducted both theoretical and empirical analyses of how IL, teacher collaboration, and collective efficacy beliefs support student learning.

Document bibliographic coupling analysis

In bibliographic coupling analysis, all extracted data were used not only to avoid citation bias but also to identify research fronts. Using VOS-viewer to filter the 1172-node bibliographic coupling network yielded eight clusters, with 518 documents in Cluster 1, 208 in Cluster 2, 167 in Cluster 3, 152 in Cluster 4, 85 in Cluster 5, 36 in Cluster 6, 3 in Cluster 7, and 3 in Cluster 8. Figure III shows the document bibliographic coupling network, in which the node size represents the total link strength of the article. According to van Eck and Waltman (2020), a bibliographic coupling link is a link between two items that both cite the same document. The total link strength of a document is the sum of the strengths of its links with other documents.

Table II shows the number of documents in each cluster at the Foundation, Consolidation, Growth, and Maturity stages. As can be seen, in the Foundation stage, Cluster 1 had the highest number of documents (51.31%), significantly higher than the other clusters, followed by Clusters 4 (35.08%), and 3 (6.81%). In the Consolidation stage, Cluster 1 had the highest number of documents (37.60%), though less than that in the Foundation stage, followed also by Clusters 4 (25.60%) and 3

FIGURE III. Document bibliographic coupling network on IL identified by VOSviewer



Source: Compiled by author

(18.40%) with fewer publications compared with the preceding stage. In the Growth stage, Cluster 1 still had the highest number of documents (48.92%), though less than that in the Foundation stage, followed by Clusters 2 (16.02%) and 4 (11.69%), which in contrast showed increase in publications compared with the preceding stage. Finally, in the Maturity stage, the number of documents among the clusters showed bigger differences and the top three were Clusters 1 (41.60%), 2 (24.80%), and 3 (16.96%). The changing trend over the years revealed similar research focuses, evidenced by the same significant cluster (Cluster 1) in both Foundation, Consolidation, and Growth stages, but more diverse research interests in more recent years of the Maturity stage.

This study conducted content analysis of the top three publications with the largest total link strength in each cluster and identified a common theme within each cluster, as shown in Table III and discussed below.

TABLE II. Number of documents in each cluster at the four development stages, 1974-2020

Stage	Year	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
Foundation stage (1974-2000)	1974-1978	19							
	1979-1983	24			6				
	1984-1988	24			24				
	1989-1993	8			13	9			
Percentage	1994-2000	23		13	24	4			
		51.31	0	6.81	35.08	6.80	0	0	0
Consolidation stage (2001-2008)	2001-2004	20	4	7	21				
	2005-2008	27	12	16	11	7			
Percentage		37.60	12.80	18.40	25.60	5.60	0	0	0
Growth stage (2009-2013)	2009-2011	52	16	12	17	9	6		
	2012-2013	61	21	13	10	7	4	2	1
Percentage		48.92	16.02	10.82	11.69	6.93	4.33	0.87	0.42
Maturity stage (2014-present)	2014-2016	81	66	31	13	10	6	1	
	2017-2020	179	89	75	13	39	20		2
Percentage		41.60	24.80	16.96	4.16	7.84	4.16	0.16	0.32

Source: Compiled by author

TABLE III. Significant publications in each cluster of document bibliographic coupling network on IL identified by VOSviewer

Cluster	Rank	Author(s)/Year	Total Link Strength
Cluster 1 IL and professional learning communities for learning success	1	Zheng et al. (2019)	1004
	2	Schecter (2008)	656
	3	Riehl (2000)	614
Cluster 2 Impacts of shared IL practice in schools	1	Paletta et al. (2017)	1906
	2	Urick (2016)	1824
	3	Bellibas et al. (2016)	1563
Cluster 3 Leadership for teaching and learning	1	Mangin and Dunsmore (2015)	1554
	2	Spillane et al. (2004)	1143
	3	Daly et al. (2013)	1122
Cluster 4 Review studies in the context of IL	1	Neumerski (2012)	1677
	2	Qian et al. (2017)	1575
	3	Hallinger (2019)	1547
Cluster 5 Influence of IL on teachers	1	Bellibas and Liu (2018)	1535
	2	Ma and Marion (2020)	1484
	3	Urick et al. (2018)	1344
Cluster 6 Influence of IL on school performance	1	Hallinger and Hosseingholizadeh (2020)	1727
	2	Sebastian et al. (2019)	1285
	3	Louis et al. (2010)	1095
Cluster 7 Instructional coaching for teachers	1	Goldring et al. (2014)	466
	2	Devine (2013)	155
	3	Devine et al. (2013)	7
Cluster 8 Influence of IL on students' achievement	1	Boston et al. (2017)	1228
	2	Smith and Smith (2018)	894
	3	Fairman and Mackenzie (2012)	422

Note: The article marked in gray also appears in Table I

Source: Compiled by author

Cluster 1 IL and professional learning communities for learning success

The common emphasis of publications in Cluster 1 is the important impacts of IL on professional learning communities for learning success. The most significant article in this cluster, namely Zheng et al. (2019) exploring the mediating effect of professional learning communities on relationship between IL and teacher self-efficacy in the context of Mainland China. Schechter (2008) highlighted the importance of the preparatory program of principals in Israel that affects their ability to foster teachers' collective learning. The study of Riehl (2000) is distinct from the rest in its inclusive educational settings. Needless to say, the leading role in a regular school differs from that in an inclusive school with students of diverse special needs.

Cluster 2 Impacts of shared IL practice in schools

Research in Cluster 2 focused on how principals practice shared IL in schools and investigated their impact on teacher performance. Paletta et al. (2017) found that schools with higher leadership scores have greater job satisfaction and higher self-efficacy among teachers, and a better educational climate. Urick (2016) concluded that principals should have similar influence over resources, safety and facilities regardless of degree of shared IL because these tasks address foundational school needs. Bellibas et al. (2016) noted from the perspective of capacity building that system leaders have in recent years increased their investment in the preparation and professional development of school leaders.

Cluster 3 Leadership for teaching and learning

Cluster 3 represents the thoughts of scholars on the relationship of leadership with teaching and learning. Mangin and Dunsmore (2015) revealed that IL with the framing of instructional coaching as a lever for teacher instructional reform influences the enactment of coaching. Spillane et al. (2004) noted that teachers working in an IL culture perform better in teaching, instructional practice, and learning improvement, which are the most proximal causes of student achievement. Daly et al. (2013) presented a study measuring leader's network position by incoming, outgoing,

and close ties; personality traits; and leader self-efficacy after controlling for demographics.

Cluster 4 Review studies in the context of IL

Studies in Cluster 4 focused on reviewing theory and leadership research in schools' organizations. Neumerski (2012) utilized a distributed lens to examine the principal, teacher leader, and coach IL literatures. Qian et al. (2017) elaborated on three dimensions with the greatest context-specific meanings for Chinese principals, namely defining purpose and direction; nurturing positive and collaborative relationships with and among teachers; and fostering professional development to enhance teacher capacity. Hallinger (2019) reviewed theory and research on educational leadership and management. The patterns thus obtained revealed that the research front in the emerging-region literature in educational leadership and management lies in papers that examine principal and shared leadership in relation to student achievement and curriculum reform.

Cluster 5 Influence of IL on teachers

Data gathering and analysis in the studies of Cluster 5 are mainly through quantitative approaches using questionnaires. These studies contributed to a growing body of research evidencing a positive effect IL on teachers, such as teacher trust, teacher collegiality, teacher efficacy, and teacher instruction. Bellibas and Liu (2018) found that principals' emphasis on instructional practice and sharing leadership can play a significant role in promoting the trust, collegiality and respect among staff. Ma and Marion (2020) indicated that IL, in terms of developing a positive learning climate, directly and positively affects teacher efficacy. Urick et al. (2018) found a direct effect of IL on math instruction in the classroom and teacher participation in math professional development.

Cluster 6 Influence of IL on school performance

Studies in Cluster 6 contributed to a growing body of research evidencing a positive effect IL on school performance, such as organizational

management, focused instruction, and collegial and collaborative environment in school for teachers. Hallinger and Hosseingholizadeh (2020) highlighted that ensuring a collegial and collaborative environment for teachers is commonly articulated by successful principals as an important aspect of IL. Findings of Sebastian et al. (2019) concluded that principals view themselves as either strong or weak on IL and organizational management skills simultaneously. Louis et al. (2010) reported that teachers' professional community and the quality of classroom instruction is a mediator on the effect of IL on student achievement.

Cluster 7 Instructional coaching for teachers

This cluster contains only three articles, published mostly during the Growth stage. Goldring et al. (2014) found that principals often experience cognitive dissonance in face of contrasting feedback from different data sources (e.g., their self-ratings to those of their teachers). Devine (2013) explored how principals' recognition of immigrant children as well as investment in supporting their learning are shaped by the logics of practice across different fields, as well as by their own authentic habitus evolving in a period of rapid social change. Devine et al. (2013) noted that instructional coaching can support schools in implementing new teaching practices in a sustained way.

Cluster 8 Influence of IL on students' achievement

This cluster also contains only three articles, published mostly during the Maturity stage. Boston et al. (2017) investigated how to support principals as instructional leaders in mathematics. Smith and Smith (2018) reported that the most impactful investment toward student achievement is helping leaders learn. The solid, sustainable, and laser-sharp focus on IL helps leaders hone, model and lead new learning through deliberate practice by engaging in rich, rigorous, and reflective open-to-learning conversations (Smith & Smith, 2018). Fairman and Mackenzie (2012) found that the work of teacher leaders results in teacher learning as well as improves students' achievement.

Examining the distribution of significant publications in the above eight clusters at the four development stages in Table II revealed the time period when the intellectual structure of studies on IL was formed. As can be seen, the majority of studies published in the Foundation stage were of Cluster 1 with contents focusing on how IL influences professional learning communities, assessing the instructional management behavior of principals, and the effect IL on school achievement. Research development further evolved from the Foundation to Consolidation stage with emphasis shifting to in-depth investigation on how IL improves teaching, and on dimensions of IL. Research in the Growth stage focused on how IL influences teaching and learning, as well as effective instructional time use for instructional leaders. Finally, in the Maturity stage, research interests become more diverse with IL explored from different perspectives, including comparing the effects of transformational and IL on student achievement. Among the wide-ranging research topics, analysis of the roles of IL, teacher collaboration, instructional strategies, and collective efficacy beliefs have received the most scholarly attention in recent years.

This study made a comparison between significant publications identified using HistCite™ tool (Table I) and VOSviewer (Table III). Of note is that the two tools yielded markedly different results. Only one significant article with high LCS, namely Neumerski (2012), was among the top 24 in the eight clusters, indicating huge discrepancy in articles identified using LCS and total link strength. The study of Neumerski (2012) was published in the Growth stage and grouped under Cluster 4, ranked sixth in LCS (Table I) and fourth in total link strength (Table III). The comparison shows that reviewing the number of times papers are cited (LCS) alone cannot objectively determine the focus of research at the development stage. In addition, the increasing number of citations from older papers over time have more “average” citations than newer papers. In contrast, bibliographic coupling is a similarity measure that uses citation analysis to establish similarity relationships between papers, which are combined into different clusters. In other words, papers grouped in the same cluster have similar content. Therefore, cross-referencing both indicators, LCS and total link strength of publications in clusters provides a more comprehensive perspective on the research focus at different stages of development and the intellectual structure of the IL knowledge base.

Keyword co-occurrence analysis

Figure IV presents the network diagram obtained from co-occurrence analysis of the 3188 keywords in the 1172 documents screened. The results revealed 18 clusters of frontier topics in IL studies. The large number of clusters generated would imply that analysis by clusters does not have much meaning or significance. Instead, this study analyzed the keywords in terms of its number or frequency of occurrences with the threshold of appearing in a minimum of 54 publications. Table IV lists the 20 keywords that met such criterion. As can be seen and as expected, “instructional leadership” is the most frequently used keyword, appearing in 243 publications, followed by “leadership” in 174, “teachers” in 110, and “achievement” in 106 publications.

The value of Avg. Pub. Year, calculated using VOSviewer according to the weighted average concept, serves as a scalar proxy indicating how ‘new’ or ‘mature’ a particular keyword is in IL research. As shown in Table IV, keywords with frequent occurrences appear in documents of earlier publication years and vice versa. Hence, the top four frequent-occurring keywords “instructional leadership”, “leadership”, “teachers”,

FIGURE IV. Network visualization of keywords co-occurrence analysis by VOSviewer

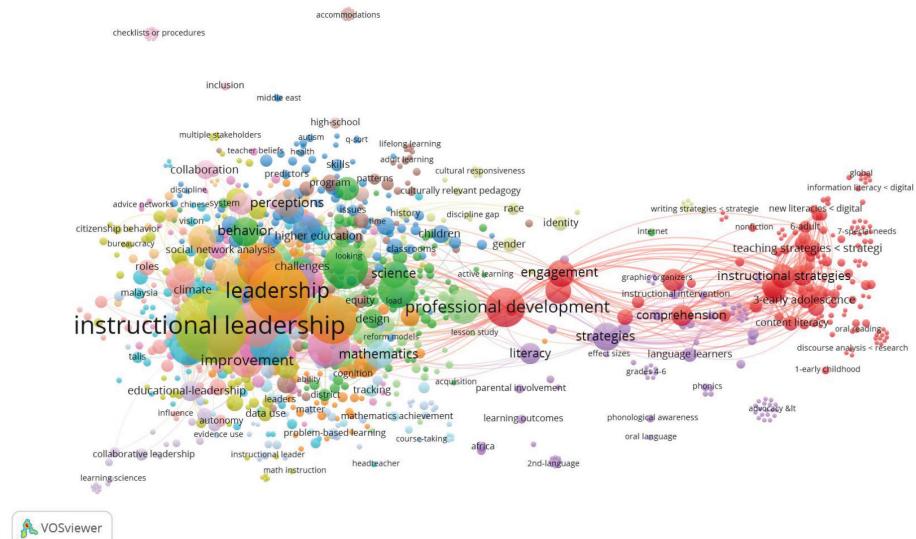


TABLE IV. Significant keywords from co-occurrence analysis

Keyword	Occurrences	Avg. Pub. Year
instructional leadership	243	2015.19
leadership	174	2014.35
teachers	110	2014.40
achievement	106	2014.53
performance	99	2014.56
education	91	2014.59
policy	85	2014.94
principals	77	2014.83
principal leadership	73	2016.26
professional development	73	2015.22
reform	70	2013.66
shared leadership	66	2016.43
higher education	65	2014.78
students	62	2016.78
school	58	2014.43
distributed leadership	57	2016.29
instruction	56	2015.32
job satisfaction	56	2015.65
improvement	54	2015.73
management	54	2014.50

Note: Occurrences refers to the number of publications the keyword appears in; Avg. Pub. Year the average publication year for articles that include the keyword

Source: Compiled by author

and “achievement” are more mature topics appearing in publications of 2014 and 2015 while keywords with low occurrences including “principal leadership”, “shared leadership”, “students”, “distributed leadership”, and “student-achievement” represent emergent research fronts mentioned in more recent publications of 2016.

Burst detection analysis

This study conducted a burst detection analysis using the Sci2 Tool to distinguish between topics of sustained research interest over time and topics that are popular merely for a few years. The burst detection

analysis identifies keywords with high-concentration and high-density characteristics in the document according to the density of changes in keyword frequency. In this way, the sudden growth of a research field can be detected in terms of the frequency with which the term subject is used. In addition, changes in research trends can be determined according to the burst weight, start and end year of each burst keyword. In this study, the top 30 keywords with the largest burst weights in IL literature were included in the analysis, the results are shown in Table V.

In the Foundation stage, there are six burst keywords. “Leadership” had the highest burst weight, followed by “school reform” and “urban schools”. They highlighted that IL practices have an effect on the school reform movement to improve student performance, especially students from poor families studying in urban schools (Polite et al., 1997). “Problem-based learning” remained in vogue for the longest duration of 19 years from 1995 to 2013 while “professional community”, though important, represented a younger focus of research and had the shortest burst of seven years from 1998 to 2004. Moreover, these burst results echoed the emergence of IL as a leadership style and model for effecting problem-based learning, professional community, and school performance (Figueroa et al., 2020; Irby, 1996).

In the Consolidation stage, there are eight burst keywords. As expected, “instructional leadership” had the highest burst weight, followed by “instructional improvement” and “professional development”. These burst results echoed that IL is a key influence on the teacher’s instructional improvement and professional development (Reitzug et al., 2008). Burst keywords including “principal leadership”, “curriculum development”, and “leadership qualities”, highlight the focus of the main research trend at this stage while other leadership styles, such as “curriculum leadership” and “teacher leadership”, have also become hot issues in IL studies on implementing curriculum reform (Hsiao et al., 2008).

In the Growth stage, there are eight burst keywords. “Higher education” had the highest burst weight, followed by “teaching styles” and “transformational leadership”. The high burst weight of “higher education” reflected the sharp increase in research on IL practice and development in universities, while that of “teaching styles” and “transformational leadership” indicated such a goal as a focal point in this stage. Despite of their significance, they appear only for a short period of time. Moreover, these burst results highlight that IL has influence on instructional coaching and school improvement (Ruebling et al., 2004). Keywords “inclusive

TABLE V. Significant keywords with highest burst weights clustered by development stage

Stage	Keywords	Weight	Start	End	1974-2020
Foundation stage (1974-2000)	leadership	2.33	1997	2000	
	school reform	2.25	1999	2010	
	urban schools	1.34	1998	2000	
	problem-based learning	1.20	1995	2013	
	professional community	1.18	1998	2004	
	school performance	0.96	2000	2009	
Consolidation stage (2001-2008)	instructional leadership	4.09	2003	2007	
	instructional improvement	2.76	2003	2012	
	professional development	1.74	2005	2009	
	principal leadership	1.63	2007	2008	
	curriculum development	1.55	2007	2011	
	leadership qualities	1.28	2007	2009	
	curriculum leadership	1.15	2008	2014	
	teacher leadership	1.10	2007	2010	
Growth stage (2009-2013)	higher education	1.68	2010	2014	
	teaching styles	1.25	2013	2013	
	transformational leadership	1.25	2013	2013	
	instructional coaching	1.21	2013	2016	
	school improvement	1.16	2012	2013	
	inclusive education	1.09	2009	2012	
	distributed leadership	1.02	2009	2010	
	educational administration	0.92	2011	2013	

(continued)

TABLE V. Significant keywords with highest burst weights clustered by development stage
(continued)

Stage	Keywords	Weight	Start	End	1974-2020
Maturity stage (2014-present)	job satisfaction	1.90	2015	2017	
	social network analysis	1.80	2018	2019	
	shared leadership	1.58	2019	2020	
	teaching strategies	1.58	2018	2020	
	systematic review	1.44	2019	2020	
	principal preparation	1.34	2019	2020	
	school climate	1.22	2019	2020	
	educational reform	1.19	2016	2017	

Source: Compiled by author

education" also reflected the sharp increase in research on IL practice and development in inclusive schools (Ruairc et al., 2012). In this stage, "distributed leadership" and "transformational leadership" received greater attention and were compared with IL in terms of effectiveness in improving school performance (Halverson & Clifford, 2013). Other notable research fronts including "educational administration" reflected the stress in recent literature on excellence in schools and the positive effect principals can have on quality instruction; thus IL has received renewed emphasis in writings on school administration (Lee & Hallinger, 2012).

The Maturity stage of IL studies had two keywords with high burst weights, namely "job satisfaction" and "social network analysis". The highest burst weight of "job satisfaction" revealed the emphasis that IL is an antecedent for job satisfaction (Skaalvik, 2020). These burst results also highlighted the usage of "social network analysis" to understand the influence of principals' social networks and how principals navigate instructional development initiatives (Rigby, 2016). Of note is that keywords such as "shared leadership", "teaching strategies", "systematic review", "principal preparation", and "school climate" have all experienced strong and recent bursts that persist till present.

While both keywords co-occurrence and burst detection analyses can be employed to explore the research fronts, they yielded different results. Take the keyword “shared leadership” for example. Co-occurrence analysis showed its Avg. Pub. Year being 2016.43 (Table IV), implying that it is a relatively new research topic most likely to be featured in recent literature. However, burst detection analysis revealed its burst starting but also ending in 2020 (Table V). Another keyword “teachers” is also an emergent research area with Avg. Pub. Year being 2014.40 (Table IV) but it is not identified by burst detection analysis (Table V). Other up-and-coming research fronts including “teaching strategies”, “systematic review”, “principal preparation”, and “school climate” had strong and recent bursts in the Maturity stage persisting till present (Table V). Despite being hot, they appear in few publications and are not listed among the significant keywords from co-occurrence analysis (Table IV).

Conclusions, implications, limitations, and suggestions for future research

This study applied science mapping methods using HistCiteTM, VOSviewer and Sci2 Tool to identify, visualize and describe the knowledge base of IL research. Four development stages, namely Foundation, Consolidation, Growth and Maturity stages were identified along with the most influential studies in each stage. Document bibliographic coupling and content analysis conducted revealed not only the knowledge base but also the intellectual structure of IL studies in each development stage. Keywords co-occurrence and burst detection analyses showed “shared leadership” as the recent focus in the field. As suggested by burst detection analysis, keywords including “teaching strategies”, “systematic review”, “principal preparation”, and “school climate” indicated also emergent fronts. Analysis results obtained using different tools were compared. The discrepancies in analysis results highlight the need for diverse analytical tools to be adopted as they complement each other in offering multiple and complementary perspectives for an across-the-board overview of IL research in the past five decades.

The implications of this study are as follows. First, IL research is still growing. Over the last 50 years, IL has continued to develop and has been reorganized, especially conceptually, demonstrating the continued

relevance of IL, both in theory and in principal leadership practice. Second, findings of this study highlight IL trends in comparing against and integration with other leadership models. In other words, there has been continuous efforts devoted to developing ideal educational leadership within the scope of schools for principals and school organizations to apply according to the respective school context. Third, the knowledge base of IL has evolved for almost five decades and remains to be an intellectual pillar for research on principal leadership. The schools of thought underlying the conceptual foundation of IL today reflect a common theme centered on how principals as instructional leaders promote student learning, teacher teaching performance, and school improvement.

This study also has limitations. First, findings in this review are obtained from an analysis of WoS-indexed bibliographic data; thus, review in this paper is only limited to assessing the evolution of the corpus of WoS-indexed publications. Second, bibliometric analyses tend to emphasize only the dominant trends of the literature. Non-dominant features that may have significant potential may have been overlooked. Overcoming this deficiency would give a more comprehensive literature review on IL.

Possible directions for further research include the following. First, the adopted interdisciplinary approach to the analysis of IL research which allowed identification of new research trends, can be extended by including investigation of other bibliometric databases (such as Scopus, ERIC, and EBSCO). Second, expanding the analysis to include, e.g., co-citation and co-authorship relationships, or full-text analysis of papers, would also allow comparison of the results obtained to date. Third, analyses made using other methods or bibliometric programs (such as CiteSpace, Pajek, and SciMAT) may yield interesting results.

Bibliographical references

Barreiro, E.W. (2015). Using HistCite software to identify significant articles in subject searches of the Web of Science. *Computer Science*, 30, 45-64.

Bellibas, M.S., Bulut, O., Hallinger, P., & Wang, W.C. (2016). Developing a validated instructional leadership profile of Turkish primary school principals. *International Journal of Educational Research*, 75, 115-133.

Bellibas, M.S., & Liu, Y. (2018). The effects of principals' perceived instructional and distributed leadership practices on their perceptions of school climate. *International Journal of Leadership in Education*, 21(2), 226-244.

Blase, J., & Blase, J. (1999). Principals' instructional leadership and teacher development: Teachers' perspectives. *Educational Administration Quarterly*, 35(3), 349-378.

Boston, M.D., Henrick, E.C., Gibbons, L.K., Berebitsky, D., & Colby, G.T. (2017). Investigating how to support principals as instructional leaders in mathematics. *Journal of Research on Leadership Education*, 12(3), 183-214.

Cale, M.C., Delpino, C., & Myran, S. (2015). Instructional leadership for special education in small to mid-size urban school districts. *Advances in Educational Administration*, 22, 155-172.

Daly, A.J., Liou, Y.H., Tran, N.A., Cornelissen, F., & Park, V. (2013). The rise of neurotics: Social networks, leadership, and efficacy in district reform. *Educational Administration Quarterly*, 50(10), 1-46.

Day, C., Gu, Q., & Sammons, P. (2016). The impact of leadership on student outcomes: How successful school leaders use transformational and instructional strategies to make a difference. *Educational Administration Quarterly*, 52(2), 221-258.

Devine, D. (2013). Practising leadership in newly multi-ethnic schools: Tensions in the field? *British Journal of Sociology of Education*, 34(3), 392-411.

Devine, M., Houssemand, C., & Meyers, R. (2013). Instructional coaching for teachers: A strategy to implement new practices in the classrooms. *Procedia - Social and Behavioral Sciences*, 93, 1126-1130.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-18.

Fairman, J.C., & Mackenzie, S.V. (2012). Spheres of teacher leadership action for learning. *Professional Development in Education*, 38(2), 229-246.

Figueroa, M.C., García, J.M.T., Arós, C.B., & García, P.I. (2020). Leadership and school success in disadvantaged contexts: The principals' perspective. *Revista de Educacion*, 388, 163-187.

Gläzel, W., & Czerwon, H.J. (1996). A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*, 37(2), 195-221.

Goddard, R., Goddard, Y., Kim, E.S., & Miller, R. (2015). A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *American Journal of Education*, 121(4), 501-530.

Goldring, E.B., Mavrogordato, M., & Haynes, K.T. (2014). Multisource principal evaluation data: Principals' orientations and reactions to teacher feedback regarding their leadership effectiveness. *Educational Administration Quarterly*, 51(4), 572-599.

Grissom, J.A., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher*, 42(8), 433-444.

Hallinger, P. (2019). Science mapping the knowledge base on educational leadership and management from the emerging regions of Asia, Africa and Latin America, 1965–2018. *Educational Management Administration & Leadership*, 48(2), 209-230.

Hallinger, P., & Hosseingholizadeh, R. (2020). Exploring instructional leadership in Iran: A mixed methods study of high- and low-performing principals. *Educational Management Administration & Leadership*, 48(4), 595-616.

Hallinger, P., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, 86(2), 217-247.

Hallinger, P., Wang, W.C., Chen, C.W., & Li, D. (2015). *Assessing instructional leadership with the principal instructional management rating scale*. Springer.

Halverson, R., & Clifford, M. (2013). Distributed instructional leadership in high schools. *Journal of School Leadership*, 23(2), 389-419.

Heck, R.H., Larsen, T.J., & Marcoulides, G.A. (1990). Instructional leadership and school achievement: Validation of a causal model. *Educational Administration Quarterly*, 26(2), 94-125.

Hsiao, H.C., Chen, M.N., & Yang, H.S. (2008). Leadership of vocational high school principals in curriculum reform: A case study in Taiwan. *International Journal of Educational Development*, 28, 669-686.

Hsieh, H.F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.

Irby, D.M. (1996). Models of faculty development for problem-based learning. *Advances in Health Sciences Education*, 1(1), 69-81.

Juma, J.J., Ndwiga, Z.N., & Nyaga, M. (2021). Instructional leadership as a controlling function in secondary schools in Rangwe Sub County,

Kenya: Influence on students' learning outcomes. *Educational Management Administration & Leadership*, 1-18.

Kleinberg, J. (2003). Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, 7, 373-397.

Lee, M., & Hallinger, P. (2012). National contexts influencing principals' time use and allocation: Economic development, societal culture, and educational system. *School Effectiveness and School Improvement*, 23(4), 461-482.

Leithwood, K., & Duke, D.L. (2009). Mapping the conceptual terrain of leadership: A critical point of departure for cross-cultural studies. *Peabody Journal of Education*, 73(2), 31-50.

Li, K., Rollins, J. & Yan, E. (2018). Web of Science use in published research and review papers 1997-2017: A selective, dynamic, cross-domain, content-based analysis. *Scientometrics*, 115, 1-20.

Louis, K.S., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, 21(3), 315-336.

Ma, X., & Marion, R. (2020). Exploring how instructional leadership affects teacher efficacy: A multilevel analysis. *Educational Management Administration & Leadership*, 49(1), 188-207.

Mangin, M.M., & Dunsmore, K.L. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly*, 51(2), 179-213.

Marks, H.M., & Printy, S.M. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 39(3), 370-397.

Meza, C.S.R. (2021). A bibliometric analysis of academic literacy: A review of the state of the art, from the past to the future. *Revista de Educacion*, 394, 63-94.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269.

Murphy, J., Elliott, S.N., Goldring, E., & Porter, A.C. (2007). Leadership for learning: A research-based model and taxonomy of behaviors. *School Leadership & Management*, 27(2), 179-201.

Neumerski, C.M. (2012). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional

leadership, and where should we go from here? *Educational Administration Quarterly*, 49(2), 310-347.

Paletta, A., Alivernini, F., & Manganelli, S. (2017). Leadership for learning: The relationships between school context, principal leadership and mediating variables. *International Journal of Educational Management*, 31(2), 98-117.

Polite, V.C., McClure, R., & Rollie, D.L. (1997). The emerging reflective urban principal: The role of shadowing encounters. *Urban Education*, 31(5), 466-489.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 348-349.

Qian, H., Walker, A., & Li, X. (2017). The west wind vs the east wind: Instructional leadership model in China. *Journal of Educational Administration*, 55(2), 186-206.

Radhakrishnan, S., Erbis, S., Isaacs, J.A., & Kamarthi, S. (2017). Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLOS ONE*, 12(3), 1-16.

Reitzug, U.C., West, D.L., & Angel, R. (2008). Conceptualizing instructional leadership: The voices of principals. *Education and Urban Society*, 40(6), 694-714.

Riehl, C.J. (2000). The principal's role in creating inclusive schools for diverse students: A review of normative, empirical, and critical literature on the practice of educational administration. *Review of Educational Research*, 70(1), 55-81.

Rigby, J.G. (2013). Three logics of instructional leadership. *Educational Administration Quarterly*, 50(4), 610-644.

Rigby, J.G. (2016). Principals' conceptions of instructional leadership and their informal social networks: An exploration of the mechanisms of the mesolevel. *American Journal of Education*, 122(3), 433-464.

Robinson, V.M.J., Lloyd, C.A., & Rowe, K.J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44(5), 635-674.

Ruairc, G.M., Ottesen, E., & Precey, R. (2012). *Leadership for inclusive education: Values, vision and voices*. Sense Publishers.

Ruebling, C.E., Stow, S.B., Kayona, F.A., & Clarke, N.A. (2004). Instructional leadership: An essential ingredient for improving student learning. *Educational Forum*, 68(3), 243-253.

Schecter, C. (2008). Exploring success-based learning as an instructional framework in principal preparatory programs. *Journal of School Leadership*, 18(1), 62-95.

Sci2 Team. (2009). *Science of science (Sci2) tool*. Indiana University and SciTech Strategies.

Sebastian, J., Allensworth, E., Wiedermann, W., Hochbein, C., & Cunningham, M. (2019). Principal leadership and school performance: An examination of instructional leadership and organizational management. *Leadership and Policy in Schools*, 18(4), 591-613.

Shatzer, R.H., Caldarella, P., Hallam, P.R., & Brown, B.L. (2014). Comparing the effects of instructional and transformational leadership on student achievement: Implications for practice. *Educational Management Administration & Leadership*, 42(4), 445-459.

Singh, V.K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics*, 126, 5113-5142.

Skaalvik, C. (2020). Self-efficacy for instructional leadership: Relations with perceived job demands and job resources, emotional exhaustion, job satisfaction, and motivation to quit. *Social Psychology of Education*, 23, 1343-1366.

Smith, J.R., & Smith, R.L. (2018). *Impact coaching: Scaling instructional leadership*. Corwin.

Spillane, J.P., Halverson, R., & Diamond, J.B. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies*, 36(1), 3-34.

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, 46(1), 31-56.

Townsend, T. (2019). *Instructional leadership and leadership for learning in schools: Understanding theories of leading*. Palgrave Macmillan.

Urwick, A. (2016). Examining US principal perception of multiple leadership styles used to practice shared instructional leadership. *Journal of Educational Administration*, 54(2), 152-172.

Urwick, A., Wilson, A.S.P., Ford, T.G., Frick, W.C., & Wronowski, M.L. (2018). Testing a framework of math progress indicators for ESSA: How opportunity to learn and instructional leadership matter. *Educational Administration Quarterly*, 54(3), 396-438.

van Eck, N.J., & Waltman, L. (2020). *VOSviewer manual*. Universiteit Leiden.

Wahlstrom, K.L., & Louis, K.S. (2008). How teachers experience principal leadership: The roles of professional community, trust, efficacy, and shared responsibility. *Educational Administration Quarterly*, 44(4), 458-495.

Weber, J.R. (1996). *Leading the instructional program*. ERIC.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research*, 74, 255-316.

Zheng, X., Yin, H., & Li, Z. (2019). Exploring the relationships among instructional leadership, professional learning communities and teacher self-efficacy in China. *Educational Management Administration & Leadership*, 47(6), 843-859.

Zyoud, S.H., Waring, W.S., Al-Jabi, S.W., & Sweileh, W.M. (2017). Global cocaine intoxication research trends during 1975-2015: A bibliometric analysis of Web of Science publications. *Substance Abuse Treatment, Prevention, and Policy*, 12(6), 1-16.

Contact address: Imam Gunawan, National Tsing Hua University, International Intercollegiate Ph.D. Program, No. 101, Section 2, Kuang-Fu Road, Hsinchu 300044, Taiwan R.O.C.; Universitas Negeri Malang, Department of Educational Administration, Jl. Semarang No. 5, Malang 65145, Indonesia. E-mail: imam.gunawan.fip@um.ac.id



Reviews

Desmuget, M., (2020). *The Factory of Digital Cretins. The dangers of screens for our children.* Barcelona: Península. 442 pp. ISBN: 978-84-9942-933-5.

Just two years ago, Michel Desmurget—a prestigious French neuroscientist and director of research at the National Institute for Health and Medical Research in France— published *The Factory of Digital Cretins. The dangers of screens for our children* — Femina award for French letters —, a work in which she addresses digital consumption and its consequences on childhood and adolescence.

The study - which is divided into two parts with several chapters and an epilogue - addresses the numerous harmful effects that the use of technology has on children and adolescents in full intellectual development through examples, graphs, comparisons, analysis of previous studies and various investigations.

At first, the author shows the effects that the abusive use of mobile devices from an early age has on the hippocampus and the cerebellum of young people, as shown by the numerous and rigorous scientific studies included in the book and that alarm about the disastrous consequences. that it carries in their health and their learning. Sequels that are even greater if the adolescent comes from a disadvantaged family since the consumption time is greater and the parental control is less.

Likewise, Desmurget points out that the population is not aware of these risks because contradictory opinions are expressed in the media that come from specialists but also from people not versed in the subject who defend the interests of large technology companies before social interests. It supposes —according to the specialist— a significant difficulty for the citizen to determine which sources are reliable and to synthesize all the information that is broadcast daily in the media in which unreliable data, fallacious arguments or unproven evidence are mixed with scientific evidence. which causes a flood of data impossible to process.

The author demonstrates that there are no digital natives and that technology does not improve any brain at all according to the scientific evidence available to date. In addition, the skills that can be learned in a digital medium are not inferred —for the most part— to other areas, hence, as Desmurget exemplifies, if a child learns to play a video game, he only learns that, since he cannot extrapolate that learning in any area of your life.

The second part of the work is intended to review the consequences that these users suffer in their physiological system after the early and continued use of technological means. The duration or efficiency of sleep and the memorization of what has been learned are altered, which has disastrous repercussions in the school environment and their learning. The research shows that the scientific evidence ensures that the use of screens —whether for recreational or educational purposes— causes lower school performance and an impoverishment in the health of students, for this reason it is essential that minors do not develop a dependency on technology from an early age to avoid later problems.

For all these reasons, the neuroscientist assures that the work and control of families who must avoid the use of technology before the age of six is essential to promote good cognitive development and good academic results. It is devastating to verify that the continued use of these devices at an early age affects human interaction, the child's language and their ability to concentrate, attacking the basic pillars of human identity.

Desmurget's study demonstrates —in conclusion— that screens affect language, concentration, memory and learning in children and adolescents, pointing out that all of this is a public health problem that governments must address and that society you should worry.

A work that should not go unnoticed by anyone if we want to have a full development and live in a full society.

María de las Nieves García Pareja

Serrano, J.L. (2022). Computational Thinking in Education, by José Luis Serrano. 90 pp. ISBN: 978-8409450435

The book is a perfect synthesis of everything you need to know about computational thinking (hereafter, CT) in education, so that teachers and educators can enter this world without making mistakes. It provides the reader with a good conceptual basis and a set of useful tools to make reasoned, critical and well-founded decisions on how, when and why to encourage computational thinking in students. José Luis Serrano manages to combine the results of the most recent research with his own experience and pedagogical knowledge, resulting in a unique manual that stands out for the clarity of ideas and simplicity of style despite the complexity of the subject matter.

The first chapter summarizes the long and intense road he has traveled before being able to write, as well as the motivations that have driven him to do so. He explains the serious relationships he finds between CP, emotions, problem solving and people's daily behaviors; and introduces as a novelty the 5PC Model.

In the second chapter, José Luis Serrano tries to deepen our understanding of the concept of CP, so much discussed over the years. He reviews the evolution of the concept, explaining that the current definition most commonly used is based on Wing's 2006 extension of the concept promoted by Papert. He comments that the growing tendency to consider the concept as indispensable for any person has gone hand in hand with companies such as Microsoft and Google, but has not been exempt from certain criticisms. He ends the chapter with what he considers to be the most accurate definition: "a set of cognitive processes that enables the formulation and solution of a problem to be expressed in a way that a thinking agent (human or machine) can carry out" (p. 23).

The third chapter is one of the fundamental chapters for a complete understanding of the concept of CP, since it sets out the elements that are part of it. It establishes four categories to classify the concepts associated with CP: cognitive processes, methods, practices and transversal skills. The differentiation and understanding of these allows an adequate approach to the teaching-learning of CP.

The fourth chapter shows the main theoretical references in which the didactic strategies used to develop CP are usually framed. These are: constructionism, active and experiential learning, problem-based learning and game-based learning. Each of them puts the nuance on an aspect

that the author discusses in detail. He also points to feedback as an essential learning tool, for which robots can be very useful.

The general tendency has been to identify robotics and programming with the PC, creating confusion as to what actually develops it. Chapter 5 details these issues, explaining that programming is not essential for developing PC and outlining the two main strategies for developing it: unplugged activities and robot programming.

Chapter six opens the second section of the book and tries to contextualize the 5PC model that will be presented in the following chapter. To do so, it summarizes all that has been said previously, answering fundamental questions.

In the seventh chapter, he presents the 5PC Model as "a proposal to clarify more precisely what PC is, its elements, its relationship with problem solving and the practices most commonly used for its development" (p. 54). It consists of five sequential steps to be followed, in which peripheral skills, PC components, practices used in the implementation of solutions, and problem-solving steps are activated. The chapter develops each of the points in detail, giving the keys to favor CP in teachers in training, although it is applicable to students in lower educational stages.

Finally, in the eighth chapter, the reader is provided with a collection of educational resources and examples of practices that can be very useful for taking action, as well as for resorting to the most up-to-date sources.

In short, this is a very practical manual, which is essential for teachers to be able to face the teaching-learning of PC from a critical and updated perspective.

Noemí C. Serrano García

Revista de Educación is a scientific publication of the Spanish Ministerio de Educación y Formación Profesional. Founded in 1940, with the title ‘Revista de Educación’ since 1952, it has been an exceptional witness of the evolution of Education in the last decades, as well as a regarded channel for the diffusion of the advances in Research and Innovation in the field of Education from a national and international perspective. *Revista de Educación* is Subdirección General de Atención al Ciudadano, published by the Documentación y Publicaciones, and is at present attached to the Instituto Nacional de Evaluación Educativa de la Dirección General de Evaluación y Cooperación Territorial.