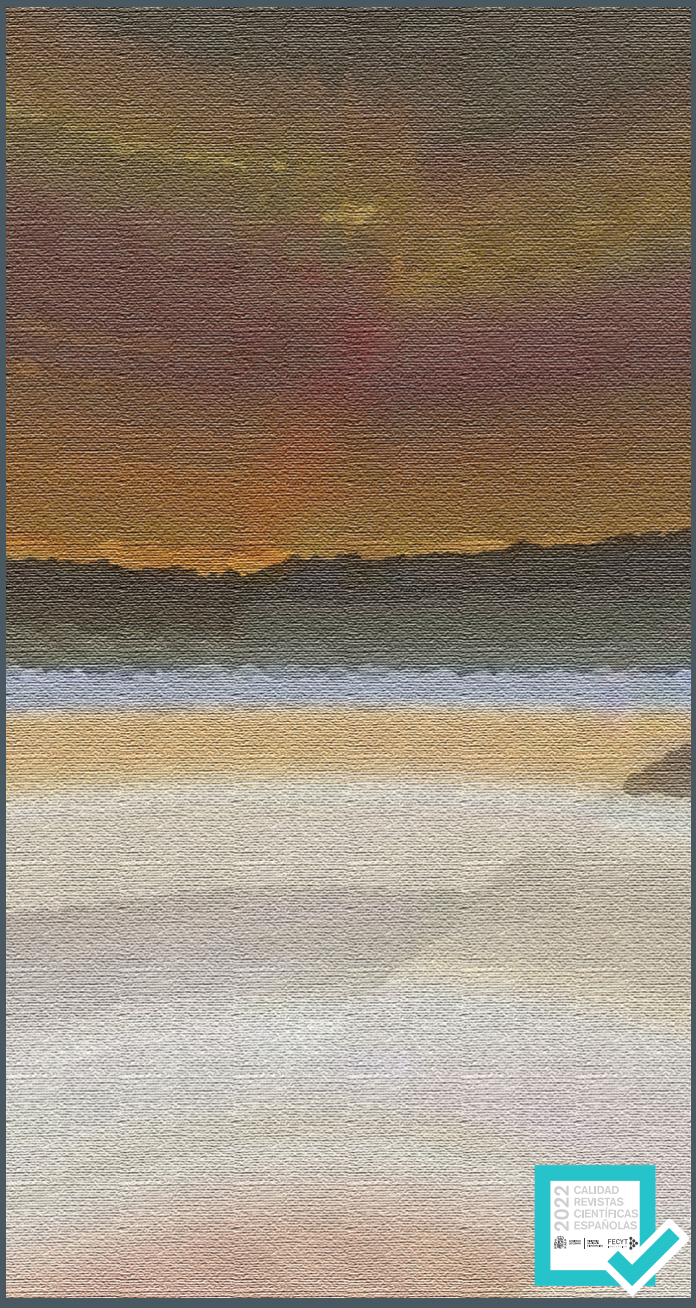
revista de EDUCACIÓN

Nº 402 OCTOBER-DECEMBER 2023



GOBIERNO DE ESPAÑA
MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL

revista de
EDUCACIÓN

Nº 402 OCTOBER-DECEMBER 2023

revista de EDUCACIÓN

Nº 402 October-December 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

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-26966-2023

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

Index

Research

Research section

Family-school participation. Evaluation from the QFIS integral model

Participación familia-escuela. Evaluación desde el modelo integral QFIS

https://doi.org/10.4438/1988-592X-RE-2023-402-593

María de los Ángeles Hernández-Prados
https://orcid.org/0000-0002-3617-215X
Universidad de Murcia

María Ángeles Gomariz Vicente
https://orcid.org/000-0002-8308-0848
Universidad de Murcia

María Paz García Sanz
https://orcid.org/0000-0003-0367-7407
Universidad de Murcia

Joaquín Parra Martínez
https://orcid.org/0000-0001-6818-8909
Universidad de Murcia

Abstract

This research presents an original instrument (Questionnaire Family Involvement in School, QFIS) to evaluate something as important as family involvement in school life. The questionnaire has been validated to measure the dimensions of such involvement: Communication with families, Participation in school activities, Sense of belonging, Home involvement, Activities in the Parent-Teacher Association (PTA) and School Council, Community involvement,

Received: 23-06-2022 Accepted: 08-02-2023

and Family training. The participants were 3612 families of students in Early Childhood Education, Primary Education, and Secondary Education in a multicultural context. The instrument was collaboratively constructed by the authors and contributions from school management teams and AMPA. After validating the content of the instrument and the corresponding exploratory factor analysis (using the principal component method), the confirmatory factor analysis (by modelling structural equations) empirically demonstrated the fit of the initial theoretical model. The calculated reliability indices were satisfactory, indicating that the resulting instrument is valid and reliable for the overall measurement of family involvement in school. The possibilities for transfer related to detecting needs for educational administrations, centre managers, and future work that allows for digital evaluation processes of involvement are evident.

Keywords: family involvement, school, structural equation model, construct validity, factor analysis.

Resumen

Esta investigación ofrece un instrumento original (Questionnaire Family Involvement in School, QFIS) para evaluar algo tan importante como es la participación de las familias en la vida de los centros escolares. Se ha validado el cuestionario para medir las dimensiones de dicha participación: Comunicación con las familias, Participación en actividades del centro, Sentimiento de pertenencia, Implicación en el hogar, Actividades en las AMPA y en el Consejo Escolar, Participación comunitaria y Formación de familias. Los participantes fueron 3612 familias de alumnado de Educación Infantil, Educación Primaria y Educación Secundaria en un contexto multicultural. El instrumento fue construido colaborativamente por los autores y las aportaciones de equipos directivos de centros y AMPA. Al validar el contenido del instrumento y el correspondiente análisis factorial exploratorio (utilizando el método de componentes principales), el análisis factorial confirmatorio (mediante modelado de ecuaciones estructurales) demostró empíricamente el ajuste del modelo teórico inicial. Los índices de confiabilidad calculados fueron satisfactorios, lo que nos informa que el instrumento resultante es válido y confiable para la medición global de la participación familiar en la escuela. Son evidentes las posibilidades de transferencia referidas a la detección de necesidades para las administraciones educativas, los gestores de centros y futuros trabajos que permitan procesos de digitalización de evaluación de la participación.

Palabras clave: participación familiar, escuela, modelo de ecuaciones estructurales, validez de constructo, análisis factorial.

Received: 23-06-2022 Accepted: 08-02-2023

Introduction

Family involvement in schools has been seen, by researchers and educational agents, as a relevant but complex process (Baker et al., 2016; Epstein et al., 2019; Kurtulmus, 2016; Wilder, 2014). Its relevance lies in a positive impact on inclusion, academic performance improvement, school climate, prevention of violent behaviour, and school dropout, among others (Jeynes, 2023; Merchán-Ríos et al., 2023; Tan et al., 2020; Wilder, 2014). Its complexity lies, on the one hand, in the interpretation made of the concept of participation, understood mostly as a face-to-face act, and not so much from an educational commitment to the school (Baker et al., 2016), reducing families to a consumer-client role, using the terminology of Vogels, which is unenterprising and un-innovative (Cárcamo y Jarpa-Arriagada, 2021). On the other hand, complexity lies in the reference to two contexts with different educational ways of proceeding, complementary and evoked to understanding (Hernández-Prados, 2022), and in the diversity of dimensions and variables that affect the family-school relationship, and its multilevel nature (Cárcamo & Jarpa-Arriagada, 2021; Epstein et al., 2019; Fernández-Alonso et al., 2017). Additionally, family involvement is subject to the geographical and cultural peculiarities of the context (Fernández-Vega & Cárcamo, 2021; Garbacz et al., 2019).

Specific research on family involvement has been characterised, firstly, by analysing the levels of such involvement from the perception of teachers, management teams, the families themselves and, on rare occasions, students (Torrego, 2019). And secondly, by employing different methodologies, information collection instruments, population sectors and statistical analyses, which make comparative analyses difficult (McNeal, 2012). According to Boonk et al. (2018), this distinction is due to research being conducted without a widely accepted theoretical framework. Hence the importance of confirmatory analyses that help us provide consistency to theoretical models.

Literature review

Recent bibliographic studies conclude that research on family involvement has mainly focused on the modalities, variables, effects, and obstacles of collaboration, as well as on the immigrant group (Egido, 2020).

Received: 23-06-2022 Accepted: 08-02-2023

Participation can be divided into both in school and at home, direct or indirect, individual or collective (Castro et al., 2015; Tan et al., 2020). Information can be gathered on the nature of parental involvement (consumer, client, participant, and partner), the level of parental involvement (informative, consultative, collaborative, decision-making, and efficacy control), and the place of participation (inside or outside the school), according to the model of Cárcamo and Jarpa-Arriagada (2021). Other aspects, such as promoting participation opportunities, improving communication, welcoming families, sharing time, and favouring the transition from involvement to commitment, are also important for family involvement (Baker et al., 2016). Without detracting from the relevance of any of them, we focus on the theoretical delimitation of each of the modalities or avenues for participation, starting from the general to the particular. In this regard, the Anglo-Saxon model of Epstein (Epstein et al., 2019), internationally recognised, remains in force, and contemplates six forms of family involvement: parenting, communication, volunteering, home learning, decision-making, and community collaboration. The North American model of Hoover-Dempsey and Sandler (Walker et al., 2005) consists of five levels of involvement (parental role beliefs, parental self-efficacy to help, feeling invited to participate by children and teachers, and participation in school activities) and recognises the importance of feeling welcomed by teachers in the decision to participate. For its part, León and Fernández (2017) present a model of nine aspects that are grouped into four modalities of family involvement: relationship (communication and family-teacher-school relationship); pedagogical support (orientation and promotion); participation (modes of participation, personal interest, knowledge); and training (behavioural and academic aspects). These studies reveal that communication, participation in school activities, family involvement, and training are common elements. The Dual Navigation Approach (DNA) model of Jeynes (2023) highlights communication, associationism, homework supervision, participation in classroom and school activities, and community resource mobilisation as aspects of school involvement, differentiating it from involvement at home.

More specifically, communication has been recognised as one of the most effective forms of student progress (Clark et al., 2019); motivating participation in the rest of the modalities, especially with school activities, home involvement, and a sense of belonging (Garreta & Llevot,

Received: 23-06-2022 Accepted: 08-02-2023

2022; Gomariz et al., 2020; Tan et al., 2020); and effective for receiving guidelines to help students at home; in fact, families demand more timely, proactive, and preventive communication (Baker et al., 2016). Communication is mainly developed through face-to-face interactions (Conus & Fahrni, 2019), often associated with attending informational meetings and tutorials to address problems (López-Castro & Pantoja, 2016; Tran, 2014), although digital communication is on the rise, despite the scepticism and distrust of teachers (Novianti & Garzia, 2020; Papakonstantinou, 2023). It includes aspects such as individual tutoring, collective meetings with families, the school agenda, informative notes, as well as any other form of communication (Consejo Escolar del Estado, 2014; Dettmers et al., 2019; Garbacz et al., 2019), although they can be specified as López-Castro and Pantoja (2016) do when focusing on family involvement in tutoring, knowledge of the functions and degree of satisfaction with the tutor and advisor, dedication and commitment of the tutor and use of the digitised communication in tutorial activities.

On the other hand, most models distinguish between Participation in school activities, focused on fostering collaborative actions within the school, and Involvement from home, which promotes support for children in homework and cognitive development (Gomariz et al., 2020; Hernández-Prados, 2022; Jeynes, 2023). The first of these encompasses attendance at meetings with teachers, help with classroom activities and participation in the running of the school as actions that make up participation in the school (Fernández-Alonso et al., 2017) as well as classroom activities, sports, coexistence, cultural activities, festivals, extracurricular outings, school services such as the library or canteen, school committees, fundraising and evaluative activities (Gomariz et al., 2020). The frequency or mode of participation (attendance, collaboration or involvement in management and decision-making) can be taken into account, as has been done on other occasions (Consejo Escolar, 2014).

By Implication at home, we understand those educational actions carried out by parents at home to promote their children's learning and allow for the general social capital of the child, encouraging family communication about school life, reading at home, participating in educational and cultural activities, instilling academic norms and expectations, supporting their learning at home and homework, parenting style, family norms, etc. (Boonk et al., 2018; Jeynes, 2023). The term does not imply presence in school, but rather a commitment to school education (Baker

et al., 2016), it promotes student achievement and well-being (Dettmers et al., 2019), and it is related to the sense of belonging and centre activities.

The training of families is included as one of the centre's activities (Consejo Escolar, 2014) and we consider it to be a specific form of participation, in line with trends that emphasise the importance of family information and training policies, betting on a training model in which families and teachers commit to learning together (Fernández-Alonso et al., 2017; Tran, 2014). Nonetheless, both dimensions, Training and Centre Activities, are deeply related (León & Fernández, 2017).

In Spain, there are two bodies representing families in schools: the Associations of Mothers and Fathers of Students (AMPA), which is an appropriate form of organisation that allows for social interaction between educational organisations and the immediate surroundings of schools (Calik et al., 2019; Merchán-Ríos et al, 2023) and promotes community participation, and School Councils. They are consultative bodies, subject to recent debate, with little prominence in educational research, which require greater teacher qualification to optimise their potential in terms of family involvement (García-Sanz et al., 2020), as currently presents low levels of involvement, especially from Roma and immigrant families (Garreta, 2016; Merchán-Ríos et al, 2023). According to Garreta and Llevot (2022), the AMPA is considered a support channel, since it consists of professionals who act as translators or mediators. New trends in research on the subject tend to identify various components of parental involvement, such as cultural, emotional, and psychological aspects, among others (Jeynes, 2023; Merchán-Ríos et al., 2023). In this sense, special educational needs, parental trust in educational support, academic expectations, sense of belonging, parental satisfaction and well-being, are aspects to consider (Tan et al., 2020), as well as the cultural barriers faced by certain immigrant or Roma communities (Garreta & Llevot, 2022; Merchán-Ríos et al., 2023). New models also emerge, such as the one proposed by Garbacz et al. (2019), which contemplates: communication between home and school, home expectations and monitoring, educational support, school and community participation, and school attendance. Subsequent studies have evolved by introducing new elements such as the sense of belonging, family training, and the facilitating role of the teacher, among others (Gomariz et al., 2022; Hernández-Prados et al., 2019; Hernández-Prados et al., 2015). From a more social approach to education, community participation is

Received: 23-06-2022 Accepted: 08-02-2023

considered. It involves ending the barriers around the school to collaborate with the community, through actions that promote cooperation between schools, families, organisations, community groups, businesses, and agencies (Gomariz et al., 2020). There are community participation initiatives through service-learning, which emphasise the commitment to solidarity in students and facilitate responsible citizenship (Rabadán et al., 2022), but they generally do not integrate family participation. Gahwaji (2019) includes, among others, national and religious events, celebrations, programs of local organisations, comprehensive family service, guides to community institutions, partnerships with libraries, parks and museums, and family association in collaboration with the community. From another approach, more focused on the participation for citizenship, solidarity-based, fundraising, ecological, religious, volunteering, neighbourhood activities could be included, among others, providing its own identity and differentiated from the rest of the dimensions in terms of content.

Finally, participation is closely linked to emotion, specifically the feeling of belonging of families to the educational centre. This feeling identifies with feeling welcomed and recognised by the educational community, so that one perceives oneself as a member of the centre (Hernández-Prados et al., 2015). This is a determining factor in family participation and in improving academic performance (Castro et al., 2015). With the review of previous studies (Reparaz et al., 2018; Uslu & Gizir, 2017), we have incorporated: identification with the educational project of the centre, trust in the teaching staff, defence of the centre's teams, feeling integrated and liberated from negative connotations towards the school, which translates into greater satisfaction and involvement with the activities organised by the centre, to the point of recommending it to other families. Feeling invited by teachers and maintaining positive communication is essential to feeling recognised, embraced and welcomed to school, increasing family participation and feeding back into the processes (Anderson & Minke, 2007). The study by Uslu and Gizir (2017) revealed that interpersonal relationships, involvement at home, and family participation in school are significant predictors of the sense of belonging.

As a result of the review conducted, a Comprehensive Model of Family Participation in Educational Centres (IMFIS) was created, which incorporates seven modalities of family participation: 1. Communication, 2. Centre activities, 3. Sense of belonging, 4. Involvement at home,

5. Parent-teacher association and school council, 6. Community participation, and 7. Family training. This model integrates traditional and emerging modalities, overcoming partial and conservative views. In addition to promoting a broad understanding of family participation, it allows for relationships to be established between each of them, since, although it is shown linearly, following the order used in the questionnaire, there are interconnections that each dimension maintains with the rest, weaving a network of interdependencies that better reflects the complexity of family participation.

Problem and objectives of the research

We pose the following research problem: how to validly and reliably evaluate family participation in their children's educational centre? Likewise, the general objective of the study was to construct a comprehensive questionnaire based on the theoretical model presented on family participation, to obtain knowledge about the dimensions that make up family participation in schools. The initial operational objectives were:

- To explore the latent variables or factors that make up the questionnaire.
- To confirm the theoretical model defined on family participation (IMFIS).
- To ensure the internal consistency of the questionnaire.

Method

A descriptive, non-experimental, cross-sectional, and confirmatory quantitative survey design was used in this research.

Participants

Out of an estimated population of 5022 families of students from 14 educational centres in Southeast Spain, where Infant, Primary, and Secondary Education is taught, all of them were invited to participate. Through

Received: 23-06-2022 Accepted: 08-02-2023

a volunteer sampling, 3639 families accepted the invitation (19.8% from Infant Education, 59.1% from Primary Education, and 20.3% from Secondary Education). However, after refining the data, a real sample of 3612 parents was obtained, achieving a confidence level of 97% and a sampling error of less than 1%.

Instrument

We started with an initial questionnaire with 19 situational questions and 88 items (with a 5-point scale, except for one dichotomous) on family participation grouped into 7 dimensions according to the IMFIS, which we named Questionnaire Family Involvement in School (QFIS). After a content validity check performed by 5 university professors (experts in the subject and research methodology), the 14 participating centres' management teams and the respective AMPA boards, the instrument retained the 19 questions about the informants' parents' situation, but the family participation items were reduced to 64, maintaining the initial 7 dimensions. These items are presented in Annex 1.

Procedure

The content validity of the questionnaire was carried out through email. Before applying the validated instruments to the informant families, they were translated into Arabic and English as necessary, since most non-Spanish origin families (except for Latin Americans) did not understand Spanish. The questionnaires were applied in a normal sanitary situation (non-pandemic), with the educational centres responsible for distributing them to families in paper format and collecting them once completed. These questionnaires were accompanied by a brief letter ensuring the confidentiality of the data and informed consent.

Data analysis

To respond to the first operational objective of the research, an exploratory factor analysis was carried out using the principal component

extraction method and the Varimax rotation method through the statistical package SPSS, version 24. Regarding the second objective, a confirmatory factor analysis was carried out using the structural equation modelling approach through the program AMOS, version 21. Finally, to respond to the third objective of the study, the reliability of QFIS was obtained through the calculation of Cronbach's Alpha coefficient and McDonald's Omega using the SPSS program. In all cases, a statistical significance level of α=.01 was considered.

Results

Exploratory factor analysis

Before proceeding with the exploratory factor analysis, in order to avoid multicollinearity problems among the QFIS items, the Spearman correlation coefficient was calculated between them. In no case did bivariate correlations exceed .85, thus, according to Kline (2005), no item had to be removed from the questionnaire validated by experts.

After checking the Kaiser-Meyer-Olkin sample adequacy measure (.955) and the statistical significance of the Barlett sphericity test (.000), which coincided with the dimensions of the theoretical model (IMFIS), 7 components were established, with an explained variance of 53.96%. In this first exploratory factor analysis, all the questionnaire items were included, except the last one of each dimension. These items were subjected to a second factor analysis because their content consists of the opposite process to what is fundamentally intended to be evaluated with the QFIS.

The first factor integrates 12 items, all belonging to the questionnaire dimension called Implication in the AMPA and in the School Council. The second factor is formed by 10 items that constitute the dimension Sense of belonging. The third factor contains the 13 items of the Involvement at Home dimension. The fourth factor includes all the items of the Participation in School Activities dimension (9), plus one belonging to the Communication with the School dimension (Q4: "I talk to the tutor in casual contacts at the educational centre"). It is an item with a low factor load for the factor it saturates (.382), with a very similar load to that obtained in the factor it should have saturated according to the IMFIS (.365). The fifth factor is formed by the 7 items of the Community Participation dimension.

Received: 23-06-2022 Accepted: 08-02-2023

The sixth factor integrates the 6 items of the Training dimension. The seventh and last factor includes 4 of the 5 items from the Communication with the School dimension, plus 2 items from the Implication in the AMPA and in the School Council of the centre (Q48: "I am, have been, or would be willing to be a member of the AMPA Board of the centre; Q56: "I am, have been, or would be willing to run as a representative of families in the School Council of the centre"). It can be understood that, to take an active part of the AMPA or the School Council of the centre, broad communication with the educational institution is necessary. In addition, the factorial loads regarding the saturations of the fifth factor are found in the first of them (.389 and .361, respectively), corresponding to the Implication in the AMPA and in the School Council of the centre dimension.

From this first exploratory factor analysis, it can be stated that the QFIS has hardly suffered any variation regarding the assignment of items to each dimension, in relation to the content validation carried out by the evaluators and the IMFIS, although the Communication with the educational centre dimension has been the one that has been most affected. Thus, the denomination of the 7 factors results as follows:

- Factor 1: Family involvement in the AMPA and the School Council, practically coinciding with dimension E of the questionnaire.
- Factor 2: Sense of belonging of families towards the educational centre, fully coinciding with dimension C of the questionnaire.
- Factor 3: Educational involvement of parents from home, fully coinciding with dimension D of the questionnaire.
- Factor 4: Family participation in activities organised by the centre, practically coinciding with dimension B of the questionnaire.
- Factor 5: Community participation of families, fully coinciding with dimension F of the questionnaire.
- Factor 6: Training of families to improve the education of their children, fully coinciding with dimension G of the questionnaire.
- Factor 7: Family communication with the educational centre. Although this factor is more closely related to dimension A of the questionnaire, it is the one that has remained weakest, probably because it is the most transversal dimension, as explained in the theoretical foundation of this contribution.

Confirmatory Factor Analysis

Regarding the second operational objective of the research, the Implication in the AMPA and in the School Council of the centre dimension was divided into two for the calculation of confirmatory factor analysis. Likewise, missing values were eliminated, and although the methodological literature on these values seems to disagree (Aguinis et al., 2013), a decision was made to disregard them, considering as such those cases in which the standardized observable variable scores exceeded the |3| score (Verdugo et al., 2008). In Figure I, according to the IMFIS, the correlation between the latent variables and the observable variables, their specific measurement error, as well as the covariance between the latent variables and also between the detected measurement errors, are graphically represented.

The model was computed using maximum likelihood method. As it is impractical to rely on multivariate normality assumptions, univariate normality was assessed by studying the skewness and kurtosis of each observed variable. For interpretation, the recommendations of Curran et al. (1996) were followed, who established limits for univariate normal behaviour at values up to |2| for skewness and up to |7| for kurtosis. This criterion was met for all observed variables.

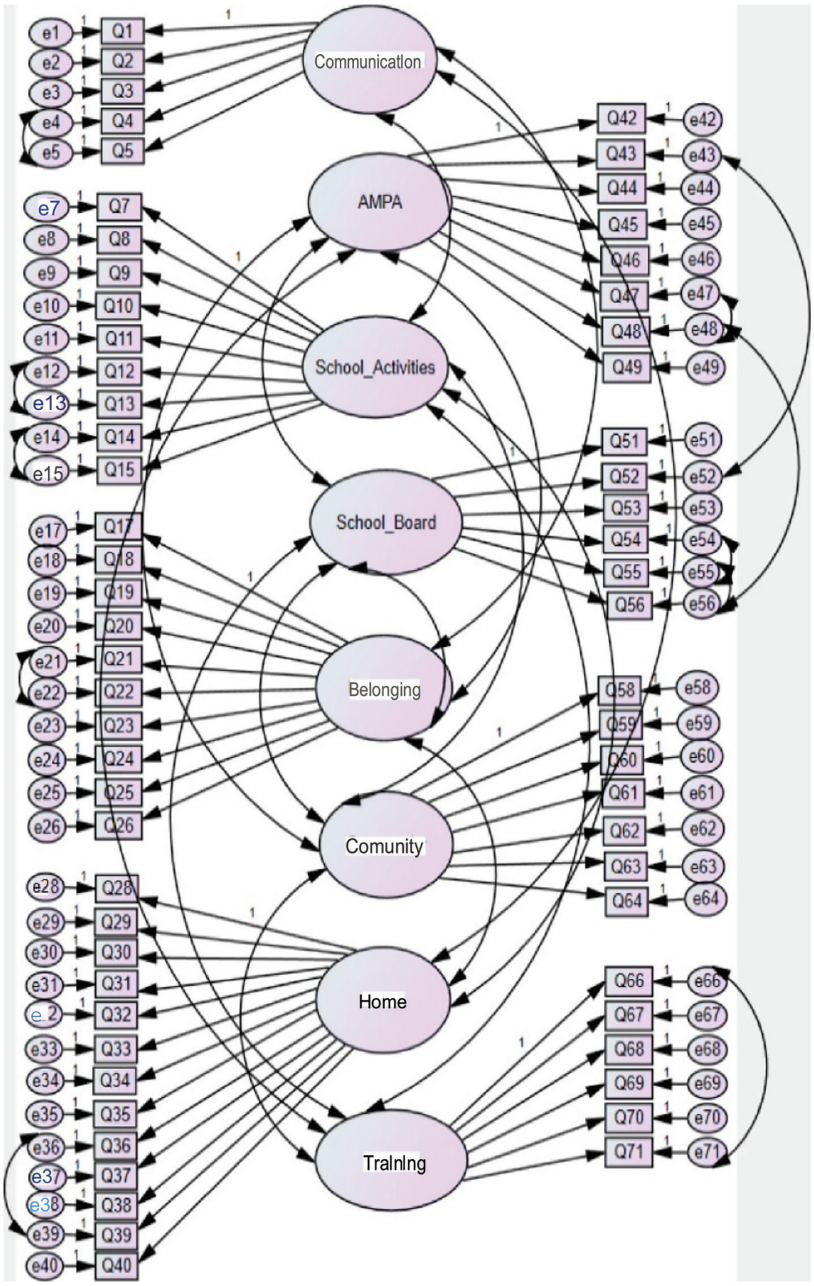
All pairs between observed and latent variables are significant, with standardized regression coefficients reaching or exceeding the value of .3 established by Cohen (1988) as the typical size of effect. Similarly, the relationships between latent variable covariance coefficients and measurement error coefficients were all significant. These correlation coefficients reach or exceed, in 88.46% of cases, the value of .3 determined by Cohen (1988).

To assess model fit, since it is recommended to use several indicators (Hu & Bentler, 1998), three indicators of different nature were used (Hair et al., 2008): normalised chi-squared or chi-squared to degrees of freedom ratio (CMIN/DF), included in parsimony goodness of fit measures; comparative fit index (CFI), integrated in incremental fit measures; and root mean square error of approximation (RMSEA), framed within absolute goodness of fit measures. Table I shows the values obtained for these indices.

Regarding the normalised chi-square, the values established by the literature range between 1 and 5 (Hair et al., 2008; Lévy & Varela, 2003; Marsh & Hocevar, 1985; Wheaton et al., 1977). The comparative fit index (CFI) should reach at least .9 (Cupani, 2012; Lévy & Varela, 2003; Marsh & Hocevar, 1985; McDonald & Marsh, 1990). The root mean square error of

Received: 23-06-2022 Accepted: 08-02-2023

FIGURE I. Questionnaire Family Involvement in School (QFIS) Structural Equation Model



Received: 23-06-2022 Accepted: 08-02-2023

Table I. The goodness-of-fit indices of the IMFIS

Index	Value
CMIN/DF	3.49
CFI	.90
RMSEA	.05

approximation (RMSEA) can be found in values lower than .08 (Browne & Cudeck, 1993; Hair et al., 2008), lower than .06 (Hu & Bentler, 1998), or scores below .05 (Lévy & Varela, 2003).

In light of the results, it can be affirmed that the IMFIS presents reasonable fit indices between the theoretical structures and the empirical data obtained, which allows us to use the questionnaire rigorously for the intended purpose.

Reliability of the questionnaire

Regarding the third objective of the study, Table 2 shows acceptable reliability indices for the questionnaire, both for the Cronbach's alpha coefficient (α) and the McDonald's omega (Ω), globally and by dimensions. All of this is in line with the categorisation established by DeVellis (2003), who determines that reliability of a measuring instrument can be considered satisfactory with values equal to or greater than .7.

TABLE II. Overall reliability of the questionnaire by dimensions.

Dimensions	α Cronbach	Ω McDonald
Global	.958	.981
Communication with the centre	.684	.661
Participation in centre activities	.848	.875
Sense of belonging	.946	.931
Home involvement	.875	.891
Participation in the AMPA and School Council	.934	.926
Community participation	.861	.880
Training	.825	.792

Received: 23-06-2022 Accepted: 08-02-2023

Discussion

The Questionnaire Family Involvement in School (QFIS) shows sufficient adjustment, although there are other instruments that have measured family involvement in school (Epstein, 2019; Garbacz et al., 2019; León & Fernández, 2017; Walker et al., 2005), its subscales allow measuring several aspects with differentiated internal consistency, integrating new dimensions that are not contemplated in other instruments or only appear as items, without considering subscales. The questionnaire presents content and construct validity with fairly explained variance. Adequate validity and reliability were observed in all dimensions, with special relevance regarding Community Participation, fully saturating the items and without considering covariance between the measurement errors of the items.

The complexity of the relationships between the dimensions of participation is consistent with what is presented in the theoretical model. More specifically, in the Communication dimension, it is observed that the item referring to casual contacts with the tutor (Q4) has saturated in the centre activities dimension when several authors consider it an informal mode of communication (Epstein, 2019), essential when parents work (Snell et al., 2020) or in immigrant families (Garreta & Llevot, 2022), decreasing such contacts as the school year progresses (Conus & Fahrni, 2019). Conversely, two items belonging to other dimensions saturate. Specifically, Q48 and Q56, referring to the willingness to belong to the AMPA board of directors and to represent families on the School Council, probably because this attitude arises from maintaining effective communication with teachers and school leaders. This type of participation has benefits for both students and the relationships between parents and teachers, increasing trust and communication (Murray et al., 2019).

Covariance has been identified between the measurement errors of two variables referring to casual contacts with the tutor (Q4) and communication with other teachers (Q5). Regarding this, we emphasise that although contacts with tutors (Kurtulmus, 2016) acting as mediators of communication between the family and non-tutor teachers, communication with the rest of the teachers should be encouraged, especially in secondary school, where the tutor only teaches one subject. The teaching functions marked by the regulations support the duty of providing "periodic information to families on the learning process of their sons and daughters, as well as guidance for their cooperation in it" (LOMLOE, Art. 91).

Although the AMPA and School Council dimensions present internal consistency independently, in the theoretical model, they constitute a single dimension. They have presented the largest volume of covariances in measurement errors, being recorded in items referring to the AMPA (Q47-Q48), the School Council (Q54-Q55-Q56), and both (Q48-Q56 and Q43-Q52). In the first covariance, we agree with Garreta (2016) that being or being willing to be a member of the AMPA Board of Directors (Q48) necessarily implies participating in the activities organised by the AMPA (Q47). Thus, the decision-making and management processes of the activities are carried out by the association's board, but all families can participate in the organised activities (Calik et al., 2019). In addition, the success of the organised activities depends on both the leadership style and the participation of families (Ndubi & Mugambi, 2019). In the measurement errors recorded in the items referring to the School Council, we can verify that the formulation between being informed of the elections to the School Council (Q54), participating in them (Q55), and presenting oneself as a representative (Q56) is similar, but the levels of participation and the responsibility they entail are different (Consejo Escolar, 2014).

Finally, the theoretical model acknowledges the interactions between the AMPA (Association of Parents of Students) and the School Council, with the latter being the great unknown (Gomariz et al., 2020). Sometimes, families play both roles: being part of the AMPA Board of Directors (Q48) and representing families on the School Council (Q56), leading to oversaturation (García-Sanz et al., 2020). For this same reason, errors related to knowing the members of the AMPA Board of Directors (Q43) and the representatives of the School Council (Q56) arise.

The dimension related to family participation in activities organised by the school is consistent, although families take advantage of these meetings to communicate informally with the tutor. The diversity of activities implies formulating items referring to categories and exemplifying the most common situations in parentheses. These are complex items (Medina, 2015), but necessary to avoid further expanding the questionnaire and exhausting the respondents. The presence of measurement errors leads us to rethink some of the items (Q12-Q13, Q14-Q15): work committees are collaboration spaces related to the communal school philosophy (Payà & Tormo, 2016), and although they have been linked to coexistence and centre improvement (Q13), they are not limited to participation in plan development but also in services (Q12). Although Stacer and Perrucci (2013)

Received: 23-06-2022 Accepted: 08-02-2023

argue that both items could constitute a single one, our proposal is to continue considering them as different. Likewise, family participation in fundraising (Q14) and centre evaluation processes (Q15) present covariance between the measurement errors of both items. There is evidence to justify the presence of both items. Some family participation taxonomies contemplate that they act as support agents to improve resource provision (Youn et al., 2012), although there are significant differences based on the social capital of families (Msila, 2012), and many impoverished communities distance themselves from school if they only care about receiving more than giving (Jeynes, 2023). Furthermore, we witness the construction of an evaluative culture based on the participation of the school community that promotes the improvement of educational quality (Janzen et al., 2017).

The dimension Involvement at home shows high consistency and reliability. It requires attention to the similarity between extracurricular or complementary activities (Q36) and cultural activities (Q37) that can lead to confusion, especially in immigrant families unfamiliar with Spanish school activities (Garreta, 2016). The difference between them lies in that the former present an academic, individual nuance, and external to the home, while the latter are typical of shared family leisure, essential in family education (Álvarez-Muñoz et al., 2023). We agree with Fernández-Alonso et al. (2017) and Castro et al. (2015) in recognising that involvement at home covers support and cultural opportunities (Q33), communication with children about school issues (Q25), and accompanying them in schoolwork (Q31).

The dimension Sense of belonging presents a level of total coincidence with the theoretical model and the highest reliability of all dimensions, with the exception of the measurement error covariance related to two items (Q21-Q22). Feeling attracted to collaborative activities or experiences with families refers to the potential of the sense of belonging as a driver for participation (Castro et al., 2015), while participating in the educational centre makes one feel part of it, emphasising how the sense of belonging to a community is generated (Dove et al., 2018; Hernández-Prados et al., 2015; Uslu & Gizir, 2017).

Finally, having information about the educational activities for families organised by the school (Q66) is the lowest level of participation in the Training dimension, since it does not imply a commitment to attend. Commitment is essential to participate, just as teacher commitment is vital to promote participation (Dove et al., 2018; Siciliano, 2016). However, when we ask if the training offered contributes to improving family-school relationships

(Q71), it is a different aspect, focused on the content of the training activity. Wilder (2014) indicates that training sessions for parents in communication with the school and involvement in home reading had positive results in family-school relationships and student performance. Given the above, there is evidence to maintain both items despite the covariance found.

Conclusions and Implications

This study offers a comprehensive model for evaluating family participation in schools and a reliable and valid multidimensional measurement tool that fits well with the IMFIS. QFIS provides relevant information for making decisions in action with families in Early Childhood, Primary, or Secondary Education centres, making it useful for different education professionals (educational administration, centre management teams, and family associations). As a limitation, although the proposed model presents an acceptable fit, it has only confirmed that it is one of the possible models (Cupani, 2012). Regarding the relationship between the obtained measurement errors, Hermida (2015) conducted 985 studies, in which 315 articles were identified that allowed correlating the measurement errors of observable variables. For this author, this is acceptable due to theoretically justifiable reasons, with the complexity of the model being a potential reason for the correlation of measurement errors. Landis et al. (2009) argue that estimating measurement errors in the structural equation model is only appropriate when these correlations are inevitable, including when observable variables share components. Both cases affect this study, with the wording of the affected items being very similar, but at the same time, all of them are necessary to conform to the specified theoretical model.

This study becomes the empirical reference framework for managing and supporting, among other transfer possibilities, the development of digital platforms for joint co-training of families and teaching staff (Ref. PID2020-113505RB-I00). We share that family empowerment in leadership roles contributes to mobilising their networks, connections, and increasing participation (Dove et al., 2018), which in turn favours the transition from a spectator family to a partner (Hernández-Prados, 2022). QFIS can promote an empathic process with teachers and the desire to share, an essential competence to establish collaborative, reciprocal family-school relationships (Peck et al., 2015). Similarly, trust between these agents

results in a more active role for families in the classroom and better support and guidance from teachers towards family education (Tran, 2014).

Funding

This work has been carried out within the framework of the ref. EDU2016-77035-R and ref. PID2020-113505RB-I00 projects, funded by MINECO and MCIN/AEI/ 10.13039/501100011033 (Ministry of Science and Innovation/State Research Agency), respectively.

Bibliographic references

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*(2), 270-301. https://doi.org/10.1177/1094428112470848

Álvarez-Muñoz, J.S., Hernández-Prados, M.A., & Belmonte, M.L. (2023). Percepción de las familias sobre los obstáculos y dificultades del ocio familiar durante el confinamiento. *Pedagogía Social. Revista Interuniversitaria, 42*, 167-180. https://doi.org/10.7179/PSRI_2023.42.11

Anderson, K. J., & Minke, K. M. (2007). Parent Involvement in Education: Toward an Understanding of Parents' Decision Making. *The Journal of Educational Research, 100*(5), 311-323. https://doi.org/10.3200/JOER.100.5.311-323

Baker, T. L., Wise, J., Kelley, G., & Skiba, R. J. (2016). Identifying barriers: Creating solutions to improve family engagement. *School Community Journal, 26*(2), 161-184. http://www.schoolcommunitynetwork.org/SCJ.aspx

Boonk, L., Gijselaers, H. J. M., Ritzen, H., & Brand-Gruwel, S. (2018). A review of the relationship between parental involvement indicators and academic achievement. *Educational Research Review, 24*, 10-30. https://doi.org/10.1016/j.edurev.2018.02.001

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen, & J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Sage.

Calik, B., Kilic, K., & Akar, H. (2019). Evaluation of the Current School-Parent Association Policy in Two Public Primary Schools. *Elementary*

Education Online, 18(1), 1-19. doi: http://ilkogretim-online.org.tr/index.php/io/article/view/2170

Cárcamo,H.,&Jarpa-Arriagada,C.(2021).Debilidadenlarelaciónfamilia-escuela, evidencias desde la mirada de futuros docentes. *Perspectiva Educacional, 60*(1), 58-80. https://doi.org/10.4151/07189729-vol.60-iss.1-art.1172

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*, 33-46. https://doi.org/10.1016/j.edurev.2015.01.002

Clark, K., Epstein, J.L., & Sanders, M.G. (2019). Measure of school, family, and community Partnerships. In J.L. Epstein (Coord.), *School, family, and community partnerships: Your handbook for action* (pp. 342-347). Corwin Press.

Cohen, J. (1988) (2ª. ed.). *Statistical Power Analysis for the Behavioral Sciences.* Academic Press.

Consejo Escolar del Estado (2014). *La participación de las familias en la educación escolar.* Ministerio de Educación, Cultura y Deporte. http://ntic.educacion.es/cee/estudioparticipacion/

Conus, X., & Fahrni, L. (2019). Routine communication between teachers and parents from minority groups: an endless misunderstanding? *Educational Review, 71*(2), 234-256. https://doi.org/10.1080/00131911.2017.1387098

Cupani, M. (2012). Análisis de ecuaciones estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación. *Revista Tesis, 1,* 186-199.

Curran, P. J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods 1*, 16-29. https://doi.org/10.1037/1082-989X.1.1.16

Dettmers, S., Yotyodying, S., & Jonkmann, K. (2019). Antecedents and outcomes of parental homework involvement: How do family-school partnerships affect parental homework involvement and student outcomes? *Frontiers in psychology, 10*, 1048. https://doi.org/10.3389/fpsyg.2019.01048

DeVellis, R.F. (2003) (2.ª ed.). *Scale development: Theory and applications*. Sage.

Dove, M. K., Zorotovich, J., & Gregg, K. (2018). School Community Connectedness and Family Participation at School. *World Journal of Education, 8(*1), 49-57. https://doi.org/10.5430/wje.v8n1p49

Received: 23-06-2022 Accepted: 08-02-2023

Egido, I. (2020). La colaboración familia-escuela: revisión de una década de literatura empírica en España (2010-2019). *Bordón: Revista de pedagogía, 72*(3), 65-84. https://doi.org/10.13042/Bordon.2020.79394

Epstein, J.L. (Coord.) (2019). *School, family, and community partnerships: Your handbook for action*. Corwin Press.

Epstein, J. L., Sanders, M. G., Sheldon, S. B., Simon, B. S., Salinas, K. C., Jansorn, N. R., Van Voorhis, F. L., Martin, C. S., Thomas, B. G., Greenfeld, M. D., Hutchins, D. J., & Williams, K. J. (2019). *School, family, and community partnerships: Your handbook for action*. Corwin Press.

Fernández-Alonso, R., Álvarez-Díaz, M., Woitschach, P., Suárez-Álvarez, J., & Cuesta, M. (2017). Parental involvement and academic performance: Less control and more communication. *Psicothema, 29*(4), 453-461. https://doi.org/10.7334/psicothema2017.181

Fernández-Vega, J. P., & Cárcamo, H. (2021). Relación familia-escuela: significados de profesores rurales sobre la participación de las familias. *Propósitos y Representaciones, 9*(2). https://doi.org/10.20511/pyr2021.v9n2.636

Gahwaji, N. M. (2019). The Implementation Of The Epstein's Model As A Partnership Framework At Saudi Kindergartens. *Journal of College Teaching & Learning (TLC), 16*(2), 11-20.

Garbacz, S. A., Hall, G. J., Young, K., Lee, Y., Youngblom, R. K., & Houlihan, D. D. (2019). Validation Study of the Family Involvement Questionnaire–Elementary Version With Families in Belize. *Assessment for Effective Intervention, 46*(3), 238–243. https://doi.org/10.1177/1534508419862857

García-Sanz, M. P., Hernández-Prados, M. Á., Galián, B., & Belmonte, M. L. (2020). Docentes, familias y órganos de representación escolar. *Estudios sobre Educación, 38*, 125-144. https://doi.org/10.15581/004.38.125-144

Garreta, J. (2016). Las asociaciones de madres y padres en los centros escolares de Cataluña: puntos fuertes y débiles. *Revista electrónica interuniversitaria de formación del profesorado, 19*(1), 47-59.

Garreta, J., & Llevot, N. (2022). Escuela y familias de origen extranjero. Canales y barreras a la comunicación en la educación primaria. *Educación XX1, 25*(2), 315-335. https://doi.org/10.5944/educxx1.31840

Gomariz, M. Á., Hernández-Prados, M. Á., Parra, J., García-Sanz, M. P., Martínez-Segura, M. J., & Galián, B. (2020). *Familias y Profesorado Compartimos Educación. Guía para educar en colaboración*. Editum. https://doi.org/10.6018/editum.2872

Gomariz, M.Á., Parra J., García-Sanz, M.P., & Hernández-Prados, M.Á. (2022). Teaching Facilitation of Family Participation in Educational Institutions. *Front. Psychol.* 12:748710. https://doi.org/10.3389/fpsyg.2021.748710

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2008) (5.ª ed.). *Análisis multivariante*. Prentice Hall.

Hermida, R. (2015). The problem of allowing correlated errors in structural equation modeling: concerns and considerations. *CMSS Computational Methods in Social Sciences, III*(1), 5-17.

Hernández-Prados, M. Á. (2022). Los ámbitos de la educación familiar: formal, no formal e informal. *Participación educativa, 12*, 29-39. https://www.educacionyfp.gob.es/mc/cee/publicaciones/revista-participacion-educativa/sumario-n12.html

Hernández-Prados, M.Á., García-Sanz, M. P., Galián, B., & Belmonte, M.L. (2019). Implicación de familias y docentes en la formación familiar. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 22*(3), 61-75. https://doi.org/10.6018/reifop.22.3.388971

Hernández-Prados, M.Á., Gomariz, M.Á., Parra, J., & García-Sanz, M.P. (2015). El sentimiento de pertenencia en la relación entre familia y escuela. *Participación educativa, 7*, 45-57. https://www.educacionyfp.gob.es/mc/cee/publicaciones/revista-participacion-educativa/sumario-n7.html

Hu, L.T., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453.

Janzen, R., Ochocka, J., Turner, L., Cook, T., Franklin, M., & Deichert, D. (2017). Building a community-based culture of evaluation. *Evaluation and Program Planning, 65,* 163-170. https://doi.org/10.1016/j.evalprogplan.2017.08.014

Jeynes, W. (2023). A Theory of Parental Involvement Based on the Results of Meta-Analyses 1. In W. Jeynes (Ed.), *Relational Aspects of Parental Involvement to Support Educational Outcomes: Parental Communication, Expectations, and Participation for Student Success* (pp. 3-21). Taylor & Francis.

Kline, R. B. (2005) (2.ª ed.). *Principle and practice of structural equation modeling*. Guilford.

Kurtulmus, Z. (2016). Analyzing Parental Involvement Dimensions in Early childhood Education. *Educational Research and Reviews, 11*(12), 1149-1153.

Landis, R., Edwards, B. D., & Cortina, J. (2009). Correlated residuals among items in the estimation of measurement models. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity, and fable in the organizational and social sciences* (pp. 195-214). Routledge.

León, V., & Fernández, M. J. (2017). Diseño y validación de un instrumento para evaluar la participación de las familias en los centros educativos. *Revista Española de Orientación y Psicopedagogía, 28*(3), 115-132. http://revistas.uned.es/index.php/reop/article/view/21622/17827

Lévy, J.P., & Varela, J. (2003). *Análisis multivariante para Ciencias Sociales*. Pearson-Prentice Hall.

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. (BOE de 30 de diciembre). https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17264

López-Castro, M., & Pantoja, A. (2016). Diseño y validación de una escala para comprobar la percepción y satisfacción de las familias andaluzas en relación con los procesos tutoriales en centros de educación primaria. *Revista Española de Orientación y Psicopedagogía, 27*(1), 47-66.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First-and higher order factor models and their invariance across groups. *Psychological Bulletin, 97*(3), 562-582. https://doi.org/10.1037/0033-2909.97.3.562

McDonald, R.P., & Marsh, H.W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin, 107*, 247-255.

McNeal, R. B. (2012). Checking in or checking out? Investigating the parent involvement reactive hypothesis. *The Journal of Educational Research, 105*, 79-89. https://doi.org/10.1080/00220671.2010.519410

Medina, F.N. (2015). Las variables complejas en investigaciones pedagógicas. *Apuntes Universitarios, 5*(2), 9-18. https://doi.org/10.17162/au.v0i2.244

Merchán-Ríos, R., Abad-Merino, S., & Segovia-Aguilar, B. (2023). Examination of the Participation of Roma Families in the Educational System: Difficulties and Successful Practices. *Multidisciplinary Journal of Educational Research, 13*(1), 64-90. https://doi.org/10.447/remie.11616

Msila, V. (2012). Black parental involvement in South African rural schools: Will parents ever help in enhancing effective school management. *Journal of Educational and Social Research, 2*(2), 303-313. https://www.richtmann.org/journal/index.php/jesr/article/view/11841

Murray B., Domina T., Renzulli L., Boylan R. (2019). Civil society goes to school: Parent-teacher associations and the equality of educational opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 5(3), 41-63. https://doi.org/10.7758/RSF.2019.5.3.03

Ndubi, P. M., & Mugambi, M. M. (2019). Factors influencing completion of parents association projects in public secondary schools in Imenti South Sub-county, Kenya. *International Academic Journal of Information Sciences and Project Management, 3*(4), 210-232. http://www.iajournals.org/articles/iajispm_v3_i4_210_232.pdf

Novianti, R., & Garzia, M. (2020). Parental engagement in children's online learning during covid-19 pandemic. *Journal of Teaching and Learning in Elementary Education (Jtlee), 3*(2), 117-131. https://doi.org/10.33578/jtlee.v3i2.7845

Papakonstantinou, A. A. (2023). Teachers' Perceptions Regarding School Parents' Online Groups. *Education Sciences, 13*(1), 60. https://doi.org/10.3390/educsci13010060

Payà, A., & Tormo, M. (2016). La participación educativa de las familias en una escuela pública valenciana. Un estudio cualitativo. *Foro de Educación, 14*(21), 227-248. https://doi.org/10.14516/fde.2016.014.021.012

Peck, N. F., Maude, S. P., & Brotherson, M. J. (2015). Understanding preschool teachers'perspectives on empathy: A qualitative inquiry. *Early Childhood Education Journal, 43*,169-179. https://doi.org/10.1007/s10643-014-0648-3

Rabadán, J. A., Benito, J., & Giorgi, A. (2022). *El aprendizaje-servicio como metodología educativa y social*. Dykinson.

Reparaz, R., Sanz, A., & González, P. (2018). *La participación de las familias en el sistema educativo de Navarra. Influencia de los factores socioeconómicos y culturales*. Gobierno de Navarra. Fondo de Publicaciones.

Siciliano, M. D. (2016). It's the quality not the quantity of ties that matters: Social networks and self-efficacy beliefs. *American Educational Research Journal, 53*(2), 227-262. https://doi.org/10.3102/0002831216629207

Snell, E. K., Hindman, A. H., & Wasik, B. A. (2020). Exploring the use of texting to support family-school engagement in early childhood settings: Teacher and family perspectives. *Early Child Development and Care, 190*(4), 447-460. https://doi.org/10.1080/03004430.2018.1479401

Received: 23-06-2022 Accepted: 08-02-2023

Stacer, M. J., & Perrucci, R. (2013). Parental Involvement with Children at School, Home, and Community. *Journal of Family and Economic Issues, 34*(3), 340–354. https://doi.org/10.1007/s10834-012-9335-y

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

Torrego, J.C. (Coord.) (2019). *La participación en los centros educativos de la Comunidad de Madrid*. Consejo Escolar de la Comunidad de Madrid. https://www.comunidad.madrid/publicacion/ref/16449

Tran, Y. (2014). Addressing reciprocity between families and schools: Why these bridges are instrumental for students' academic success. *Improving Schools, 17*(1), 18-29. http://digitalcommons.georgefox.edu/soe_faculty/99

Uslu, F., & Gizir, S. (2017). School belonging of adolescents: The role of teacher-student relationships, peer relationships and family involvement. *Educational Sciences: Theory & Practice, 17*(1), 63-82. https://doi.org/10.12738/estp.2017.1.0104

Verdugo, M.A., Crespo, M., Badía, M., & Arias, B. (2008). Metodología en la investigación sobre discapacidad. Introducción al uso de las ecuaciones estructurales. *VI Seminario Científico SAID*. Salamanca: INICO.

Walker, J. M. T., Wilkins, A. S., Dallaire, J. R., Sandler, H. M., & Hoover-Dempsey, K. V. (2005). Parental involvement: Model revision through scale development. *The Elementary School Journal, 106*, 85-105. https://www.journals.uchicago.edu/doi/abs/10.1086/499193

Wheaton, B., Muthén, B., Alwin, D.F., & Summers, G.F. (1977). Assessing relability and stability in panel models. In D.R. Heise (Ed.), *Sociological methodology* (pp. 84-136). Josseey-Bass. https://doi.org/10.2307/270754

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

Youn, M. J., Leon, J., & Lee, K. J. (2012). The influence of maternal employment on children's learning growth and the role of parental involvement, *Early Child Development and Care, 182*, 1227-1246. https://doi.org/10.1080/03004430.2011.604944

Contact address: Joaquín Parra Martínez. Universidad de Murcia, Dpto. de Métodos de Investigación y Diagnóstico en educación, Facultad de Educación. Campus de Espinardo, 30100-Murcia (España). E-mail: jparra@um.es

Received: 23-06-2022 Accepted: 08-02-2023

Annex

1. Measuring instrument

Dimensions and items of the Questionnaire Family Involvement in School, QFIS

A. Communication with the centre
1. I attend meetings with the classroom teacher.
2. I request meetings with the classroom teacher throughout the year.
3. I attend group meetings with other parents and the classroom teacher.
4. I speak with the classroom teacher in casual meetings at the beginning or end of the class.
5. I have meetings with the other teachers.
B. Involvement in school activities
6. Workshops in the classroom (Reading, handicrafts, cooking, etc.).
7. Cultural activities (historical facts, music topics, ecological, traditions, International Day of Peace, grandparents, children, women, etc.).
8. Sport activities (tournaments or displays of football, basketball, judo, karate, etc.).
9. Celebrations (Christmas, Carnival, end of year, etc.).
10. Trips (to museums, monuments, other institutions, long trips, etc.).
11. Service activities offered by the centre (homeroom, library, dining room, school transport, etc.).
12. Work commissions in the centre (Plan for Coexistence, Improvement plan, etc.).
13. Commission for classroom fundraising (gifts, costumes, classroom decorations, etc.).
14. In the processes used to assess the centre (responding to forms, using a suggestions box, making complains and/or suggestions through the AMPA (association for parents of students) or individually, etc.)
C. Sense of belonging
15. I identify with the values, ideas, attitudes, goals, etc. of the centre.
16. I consider myself to be a part of the centre.
17. If a sporting, artistic or cultural team of the centre participates in any tournament or demonstration, I support that team.
18. I trust the educational work of the teachers, supporting their decisions.
19. I find the family activities or experiences offered by the centre appealing.
20. Participating in the school makes me feel like I am a part of it.
21. Since the beginning, I've felt welcomed and integrated by the education community.
22. I'm satisfied with the education that my child receives at the school.
23. I feel free to express my ideas, concerns, suggestions, complaints, etc.
24. I would recommend this school to others with children.

Received: 23-06-2022 Accepted: 08-02-2023

D. Home involvement
25. I speak with my child about what he/she has done in class.
26. I show my child that I trust him/her.
27. I'm aware of my child's attendance.
28. I'm interested in my child's homework.
29. I'm concerned about how my child organises his/her time.
30. I promote a good study environment at home (motivating the child to study, giving him/her a good space with no distractions, learning resources, etc.).
31. I'm available to help my child with school work at any time.
32. I congratulate my child after the completing of his/her school work.
33. I give extracurricular or complementary activities to my child (languages, IT, music, dancing, sports, support lessons, etc.).
34. I promote my child's responsibility when studying, being on the lookout, but never completing the activities myself or being with the child the entire time.
35. I ensure responsible use of computers, mobile phones, etc.
36. In my family we participate in cultural activities (we read, we go to the cinema, theatre, museums, trips, concerts, exhibitions, etc.
37. I try to ensure that my child uses the things learned in class in real life
E. Involvement in the AMPA and the School Board
38. I'm aware of the structure and functioning of the AMPA.
39. I know members of the AMPA Board.
40. I'm informed of the activities organised by the AMPA.
41. I know the collection of books in which the AMPA takes part.
42. I've looked for information about the AMPA on the internet, social networks, etc.
43. I take part in activities organised by the AMPA.
44. I am, I have been or I would be willing to be a member of the AMPA Board.
45. I feel that the AMPA represents the interests of the families.
46. I'm aware of the structure of the School Board.
47. I know family representatives on the School Board.
48. I'm informed of the decisions made in School Board meetings.
49. I'm informed of the election process for the School Board (calendar, candidatures, election process, etc.)
50. I vote in the elections for the School Board.
51. I am, I have been or I would be willing to be a family representative of the School Board.
F. Community involvement
52. In collection activities (food, clothes, caps collection, charity markets, etc.).
53. In ecological activities (cleaning of rivers, demonstrations for the environment, environmental awareness programs, tree planting, etc.).

Received: 23-06-2022 Accepted: 08-02-2023

54. In neighbourhood activities (local parties, neighbour meetings, demonstrations for the needs of the neighbourhood, etc.).
55. In charity and volunteer activities (helping the elderly, the ill, those with limited resources, those who are alone, soup kitchens, etc.).
56. In activities of the different religious communities.
57. In activities targeted to diversity awareness (gender, abilities, cultural background, ethnic, etc.).
58. Activities to collaborate with youth associations to promote healthy leisure and free time activities.
G. Training
59. I'm informed of the training activities for families in the school.
60. I attend training activities for families in the school.
61. I take an active role in parent training activities (I ask questions, participate in debates, use what I've learned, etc.).
62. I'm involved in the creation of training activities for families.
63. I have sufficient training to improve my child's education.
64. The training offered by the centre helps to improve the family-school relationship.

Received: 23-06-2022 Accepted: 08-02-2023

Burnout, stress and resilience in the competitive examination process for educational corps

Burnout, estrés y resiliencia en el proceso de oposición a cuerpos educativos

https://doi.org/10.4438/1988-592X-RE-2023-402-594

Eduardo Melguizo Ibáñez
https://orcid.org/0000-0003-3693-2769
Universidad de Granada

Gabriel González Valero
https://orcid.org/0000-0001-7472-5694
Universidad de Granada

Félix Zurita Ortega
https://orcid.org/0000-0002-1189-894X
Universidad de Granada

José Luis Ortega Martín
https://orcid.org/0000-0003-4875-0188
Universidad de Granada

Abstract

Preparing for a permanent position as a teacher in the public teaching profession in Spain is a complex and tedious process that can lead to the appearance of disruptive mental states in the candidate. In light of the above, this research aims to study the levels of resilience, stress, burnout syndrome and hours of study among candidates applying for public teaching posts at different levels of education and to investigate the effects of resilience, stress and burnout syndrome on the number of hours of study spent by candidates at different levels of education through structural equation modelling. For the development of the research, a descriptive, comparative and cross-sectional study was carried out on a sample of 4117 candidates

in early childhood education, primary and secondary education. The instruments used have been validated by the scientific community and adapted to Spanish showing a high degree of reliability. In this case, the questionnaires used were the Perceived Stress Scale (PSS), the Maslach Burnout Inventory (MBI) and the Connor-Davidson Resilience Scale (CD-RISC). The results show that candidates for the pre-primary and primary education corps have higher levels of stress and burnout syndrome than candidates for the secondary education corps. In addition, secondary school teacher candidates show higher levels of resilience and a higher number of study hours. In conclusion, it is stated that the levels of burnout, stress and resilience may vary according to the educational stage for which the candidates are applying and that resilience is a fundamental element that helps to prevent and channel the negative states derived from stress and burnout syndrome.

Keywords: education system, stress, teacher burnout, resilience academic, Preschool Education, Elementary Education, Secondary Education.

Resumen

La preparación para conseguir una plaza fija como maestro del cuerpo público docente en España es un proceso complejo y tedioso que puede originar la aparición de estados mentales disruptivos en el opositor. Atendiendo a todo lo citiado anteriormente, esta investigación muestra los objetivos de estudiar los niveles de resiliencia, estrés, síndrome de burnout y horas de estudio entre los candidatos que optan a un puesto docente público en los diferentes niveles educativos e investigar los efectos de la resiliencia, el estrés y el síndrome de burnout sobre el número de horas de estudio que dedican los candidatos en los distintos niveles educativos a través de un modelo de ecuaciones estructurales. Para el desarrollo de la investigación se ha llevado a cabo un estudio descriptivo, comparativo y de corte transversal en una muestra de 4117 opositores de educación infantil, primaria y secundaria. Los instrumentos empleados han sido validados por la comunidad científica y adaptados al castellano mostrando un alto grado de fiabilidad. En este caso los cuestionarios empleados han sido la Escala de Estrés Percibido (PSS), el Inventario de Burnout de Maslach (MBI) y la Escala de Resiliencia de Connor-Davidson (CD-RISC). Atendiendo a los resultados, estos ponen de manifiesto que los opositores al cuerpo de educación infantil y primaria presentan niveles más altos de estrés y síndrome de burnout que los candidatos al cuerpo de educación secundaria. Además, los candidatos a profesores de secundaria muestran mayores niveles de resiliencia y un mayor número de horas de estudio. Como conclusion, se afirma que los niveles de burnout, estrés y resiliencia pueden variar en fucnión de la etapa educativa a la que se presentan los opositores y que la resiliencia es un elemento fundamental que ayuda a prevenir y a canalizar los estados negativos derivados del estrés y del síndrome de burnout.

Received: 19-01-2023 Accepted: 04-03-2023

Palabras clave: sistema escolar, estrés, burnout escolar, resiliencia académica, Educación Infantil, Educación Primaria, Educación Secundaria.

Introduction

Currently, the role of teachers in society is essential to understand the educational reality of every country (Parker et al., 2022). In this regard, the Spanish process of selection of teachers at the preschool, elementary, lower secondary and upper secondary levels differs from the rest of Europe in that the Spanish recruitment process involves a two-phase state examination (Suárez-Riveiro et al., 2013). The first one consists in demonstrating breath of knowledge specific to each education level that teacher candidates apply for (Real Decreto 270/2022). This demonstration is carried out through a written examination about a topic randomly chosen by the examination board where the teacher candidate sits the exam (Real Decreto 270/2022). The second phase of the examination aims to test the teaching aptitude of the candidates as well as the necessary techniques to hold a teaching position. This phase consists in the defense of a unit of work (Real Decreto 270/2022).

The process of training for these state examinations leads teacher candidates, who are commonly known as *"opositores"* in the Spanish education field, to start preparing for the public examination with high self-imposed expectations (Calderón et al., 2020). Undergoing constant levels of high expectations results in an increase in levels of stress and the onset of disruptive states that decrease the motivation needed for tasks, directly influencing the preparation for the public examination (de la Fuente y Amate, 2019).

In this regard, one of the most common states in this process is the burnout syndrome (Gutentag et al., 2017). This term is defined as a state of physical, mental and emotional exhaustion that results in a lack of interest in an activity or profession (Freudenberger, 1989; Tikkanen et al., 2022). In the education field, teachers have manifested a higher degree of physical and mental exhaustion when carrying out their duties (Agyapong et al., 2022). Likewise, the preparation for an examination over a long period of time increases the levels of burnout syndrome (Daumiller et al., 2021). It has also been observed that different variables come into play as catalyzers for the onset of physical and mental exhaustion, stress being one of those variables (Zhao et al., 2022).

Stress has been defined by numerous authors, however, this state is characterised by mental fatigue resulting from the effort to perform above the current level of competence, which contributes to the onset of mental and physical disorders (Zhen et al., 2022). Its onset is not spontaneous, but is characterised by an onset divided into three phases (Selye, 1975). The first phase consists of an alarm reaction whereby the subject becomes alert to a given stimulus (Selye, 1975). After the continuation of this phase, a resistance phase takes place, in which individuals cope with the stressor (Selye, 1975). The last phase consists of an exhaustion phase, in which the subject shows signs of mental and physical fatigue as a result of coping with the stressor stimulus (Selye, 1975).

Teaching is one of the most mentally and physically exhausting occupations, as teachers show increased levels of stress (Raducu and Stanculescu, 2022). The study conducted by de la Fuente and Amate (2019) showed that during preparation for the state examination for a public teaching post in Spain, high stress levels are the result of self-imposed high expectations as well as a sense of uncertainty. In contrast, factors such as resilience have been found to prevent increased stress levels and the occurrence of burnout syndrome (Yu et al., 2022).

Resilience is defined as the capacity that individuals possess to recover from adverse situations (Bartell et al., 2019). In the education field, resilience is defined as the aptitude to overcome any negative effect with the view to developing the competitiveness in the social, vocational and academic field despite undergoing negative effects (Mansfield et al., 2016). Resilience has been demonstrated to play a key role in the education field to prevent the onset of disruptive and unfavorable states that affect the teaching practice (Zhang et al., 2020). Specifically, the study carried out by González-Valero et al. (2021) highlights that within the Spanish educational context, it is necessary to incorporate resilience into teacher training. Similarly, Díaz-Sánchez and Barra-Almagia (2017) establish that resilience is a factor that helps to develop job satisfaction despite intrinsic adversities such as lack of rest, social obstacles and mental health. Likewise, observations have been made that while preparing to get a permanent placement in the public system, the population with the lowest levels of stress and burnout syndrome have higher levels of resilience in addition to studying more hours (Melguizo-Ibáñez et al., 2022).

In light of the above, this study addresses the following questions: Does resilience contributes to reducing the effects of stress and the

Received: 19-01-2023 Accepted: 04-03-2023

burnout syndrome? Does the education level that teacher candidates qualify for directly exert influence in terms of resilience, burnout syndrome, stress and study hours?

Finally, the main research objective is to empirically study the levels of resilience, stress, burnout syndrome and study hours among candidates for public teaching positions at different educational levels and to investigate the effects of resilience, stress and burnout syndrome on the number of study hours spent by teacher candidates at different educational levels through structural equation modelling.

Research methods and Instruments

A cross-sectional, comparative descriptive study was carried out, with the data studied at a single point in time. In this case, the data were collected from January 2022 to May of the same year.

Sample

In this case, a total of 4169 participants took part, but the final sample consisted of a total of 4117 candidates. The reduction of the sample was mainly due to the fact that some participants did not fully answer the questionnaire and because others randomly completed some questions. Looking at the gender distribution of the sample, 33.1% (n=1363) is male and 66.9% (n=2754) is female. With regard to the distribution of the stage of education being applied for, 13.9% (n=574) of teacher candidates were applying for a place in early childhood education, 76.2% (n=3134) were applying for primary education and 9.9% (n=409) were applying for secondary education.

Regarding the distribution by Autonomous Community, Table I shows a summary of the population reached by each one of them.

Instruments

- **Ad hoc sociodemographic questionnaire:** This instrument was used to study the sociodemographic variables, namely gender

Received: 19-01-2023 Accepted: 04-03-2023

Table I. Geographic distribution of the sample

	N	%
Andalusia	922	22.4%
Catalonia	158	3.8%
Community of Madrid	629	15.3%
Valencian Community	576	14.0%
Galicia	401	9.7%
Castille and Leon	254	6.2%
Basque Country	30	0.7%
Canary Islands	126	3.1%
Castille La Mancha	359	8.7%
Region of Murcia	213	5.2%
Aragon	73	1.8%
Balearic Islands	31	0.8%
Extremadura	110	2.7%
Asturias	113	2.7%
Navarre	39	0.9%
Cantabria	69	1.7%
La Rioja	14	0.3%
Total	4117	100.0%

(male/female), age, number of daily hours of study and the education level that teacher candidates were qualifying for (i.e., preschool, elementary or secondary education).

- **Perceived Stress Scale (PSS)** (Cohen et al., 1983): Due to the characteristics and Spanish descent of the sample, the adapted Spanish version of the instrument by Remor (2006) was administered. This instrument consists of 14 questions that are answered with a five-point Likert scale. In light of the level of reliability of the instrument, the reliability coefficient Cronbach's alpha was used, which showed a value of α=0.899
- **Maslach Burnout Inventory (MBI)** (Maslach & Jackson, 1981): For this study, the adapted Spanish version of the instrument

Received: 19-01-2023 Accepted: 04-03-2023

by Seisdedos (1997) was used. This instrument evaluates three dimensions of the burnout syndrome. The first dimension assesses emotional exhaustion, the second one depersonalization and the last one the level of personal accomplishment. In regard to the reliability analysis of this instrument, the coefficient Cronbach's alpha had a value of α=0.909.

- **Connor-Davidson Resilience Scale (CD-RISC)** (Connor & Davidson, 2003): Due to the geographic characteristics of the target sample, the adapted Spanish version of the instrument by Crespo et al. (2014) was used. This instrument assesses resilience using five factors for the following dimensions: persistence/tenacity/self-efficacy, control under pressure, adaptability and support networks, control and purpose, and spirituality. In this instance, the reliability analysis had a value of α=0.879.

Procedure

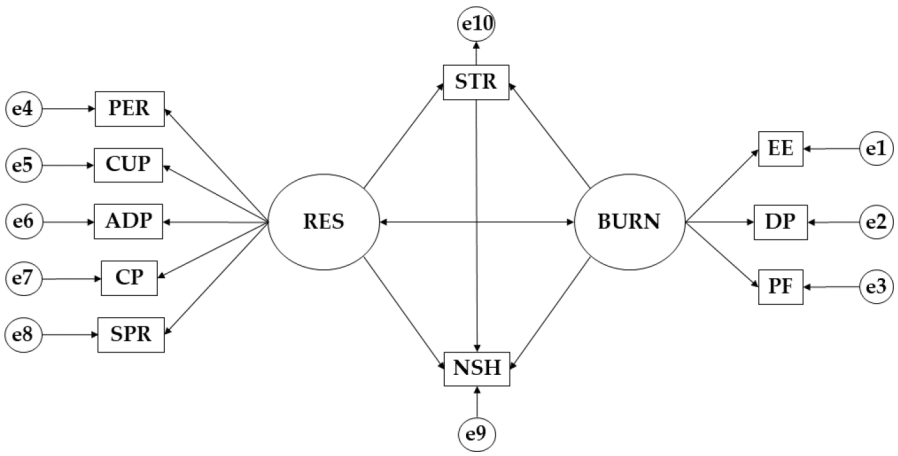
In order to collect the data, a literature study was carried out beforehand to find out about the different instruments used and their degree of reliability. Once the instruments for data collection had been established, a Google Form document was created that included these instruments and the objectives of the research. Most of the data were collected telematically. For this purpose, the different social networks were used, inviting only education opponents to collaborate. The only criterion for inclusion was that the participants had to be candidates in pre-school, primary or secondary education. In addition, all participants took part in the study on a voluntary basis. Before being given access to the instrument, they were asked for their informed consent and were assured that the data would be processed for scientific purposes and anonymously. To ensure that the responses were not random, two questionnaires were duplicated, eliminating respondents whose answers were not identical in the questionnaires. As a result, 13 respondents were discarded. Furthermore, the research followed the criteria established in the Declaration of Helsinki and was approved and supervised by an ethics committee of the University of Granada (2966/CEIH/2022).

For the comparative analysis of the results, the statistics software IBM SPSS Statics 25.0 (IBM Corp, Armonk, NY, U.S.) was used. Firstly, the

normality of the distribution of the sample was evaluated with the Kolmogorov-Sminov test. Upon obtaining a normal distribution, we proceeded to conduct an ANOVA for a single factor, examining the differences among the teacher candidates qualifying for a permanent placement in the public education system at the pre-school, elementary and secondary levels. For the analysis of the statistically significant differences, a significance level of $p \leq 0.05$. was established. To calculate the statistical power, the standardized Cohen's *d* was used (Cohen, 1992). With respect to the value obtained, the value can be classified into four levels: null (≤0.19), small (0.20-0.49), medium (0.50-0.79) and large (≥0.80).

To develop the structural equation modeling, the statistical software IBM SPSS Amos 26.0 (IBM Corp., Armonk, NY, U.S.) was used. Each proposed model consists of a total of eleven variables. Nine of them are endogenous (STR; EE; DP; PR; CP; PER; CUP; ADP; SP) and two are exogenous (BURN; RES). For the latter group of variables, a casual explanation is put forward departing from the reliability of the measures and indexes. Accordingly, the measurement errors of the distinct models were included. In respect of the direction of the arrows, unidirectional relationships are interpreted by taking the regression weights as a starting

FIGURE I. Theorethical model proposed



Note: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR); Resilience (RES); Burnout Syndrome (BURN).

Received: 19-01-2023 Accepted: 04-03-2023

point. To determine statistically significant differences, two levels of significance were established, one with a $p \leq 0.05$ and another one with a $p \leq 0.001$.

To assess the fit of the different models, the criteria established by Bentler (1990) and McDonald and Marsh (1990) were adopted. Goodness of Fit should be determined using the chi-square test, with non-significant values indicating good model fit. Also, the values of the Comparative Fit Index (CFI), Goodness of Fit Index and Incremental Fit Index should be greater than 0.900, while the Root Mean Squared Error of Approximation (RMSEA) scores should be less than 0.100 to obtain a good fit.

Results

Table II shows the results obtained from the comparative analysis. The results obtained for the stress variable indicate that the participants with the highest values are the early childhood education candidates (36.9425±7.53203). With respect to the variables related to burnout syndrome, it is obtained that the candidates for early childhood education have the highest levels of emotional exhaustion (38.5024±7.61440). Continuing with depersonalisation, it is observed that primary education candidates show the highest levels (16.6211±6.32141), while secondary education candidates had the highest levels of personalisation (26.9046±8.17550). In terms of resilience, secondary school candidates showed the highest levels of Persistence/Tenacity/Self-Efficacy (2.8020±0.75042), Adaptability and Support Networks (2.8824±0.63159), Control and Purpose (2.7330±0.74486) and Spirituality (2.4230±0.79493). In contrast, early childhood education opponents show the highest levels of Control under pressure (2.5601±0.79503).

Continuing with the structural equation modeling, the model proposed for preschool teacher candidates preparing for the state examination demonstrated good fits for each of its indexes. The chi-square test indicated a non-significant p value ($\chi 2=5.738$; df=31; pl=0.000), but in spite of the good fit, the results cannot be interpreted isolatedly given the sample size and statistic sensitivity (Tenembaum & Eklund, 2007). For this model, the values of the Comparative Fit Index (CFI), the Goodness of Fit Index (GFI) and the Incremental Fit Index (IFI) were higher than 0.900 while the RMSEA showed a value of 0.086.

TABLE II. Comparisson regarding the education level that teacher candidates were qualifying for

		N	N	S.D	F	p	ES (d)	95% CI
STR	Pre-school education	574	36.9425	7.53203	30.339	≤ 0.05	0.115[a] 0.593[b] 0.453[c]	[0.026; 0.204][a] [0.464; 0.723][b] [0.350; 0.557][c]
	Elementary education	3134	35.9588	8.76870				
	Secondary education	409	31.9438	9.53601				
EE	Pre-school education	574	38.5024	7.61440	16.844	≤ 0.05	0.470[b] 0.386[c]	[0.341; 0.598][b] [0.282; 0.489][c]
	Elementary education	3134	37.7347	7.98429				
	Secondary education	409	34.5946	9.22580				
DP	Pre-school education	574	16.1477	6.86158	18.347	> 0.05	NP	NP
	Elementary education	3134	16.6211	6.32141				
	Secondary education	409	16.1134	6.83954				
PF	Pre-school education	574	24.1735	7.58864	6.104	≤ 0.05	0.470[b] 0.608[c]	[0.341; 0.598][b] [0.284; 0.491][c]
	Elementary education	3134	23.7809	8.03754				
	Secondary education	409	26.9046	8.17550				
PER	Pre-school education	574	2.4571	0.73786	19.679	≤ 0.05	0.464[b] 0.379[c]	[0.336; 0.593][b] [0.276; 0.482][c]
	Elementary education	3134	2.5074	0.78076				
	Secondary education	409	2.8020	0.75042				
CUP	Pre-school education	574	2.5601	0.79503	12.817	≤ 0.05	0.332[b] 0.434[c]	[0.204; 0.460][b] [0.331; 0.538][c]
	Elementary education	3134	2.5533	0.87439				
	Secondary education	409	2.2726	0.95660				
ADP	Pre-school education	574	2.6317	0.62969	15.107	≤ 0.05	0.398[b] 0.339[c]	[0.270; 0.526][b] [0.236; 0.442][c]
	Elementary education	3134	2.6568	0.66976				
	Secondary education	409	2.8824	0.63159				
CP	Pre-school education	574	2.4199	0.69198	19.139	≤ 0.05	0.438[b] 0.371[c]	[0.310; 0.567][b] [0.268; 0.475][c]
	Elementary education	3134	2.4584	0.73925				
	Secondary education	409	2.7330	0.74486				
SPR	Pre-school education	574	2.3438	0.81190	2.370	≤ 0.05	0.135[c]	[0.032; 0.238][c]
	Elementary education	3134	2.3052	0.88316				
	Secondary education	409	2.4230	0.79493				
NHS	Pre-school education	574	5.0436	2.14562	12.898	≤ 0.05	0.128[c]	[0.025; 0.231][c]
	Elementary education	3134	4.9076	2.39642				
	Secondary education	409	5.2127	2.27868				

Note I: [a] Differences between pre-school and elementary education; [b] Differences between pre-school and secondary education; [c] Differences between elementary and secondary education.
Note II: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

Received: 19-01-2023 Accepted: 04-03-2023

The results obtained for preschool teacher candidates show a positive relation between stress and resilience (β=0.009). hey also indicate a positive relation between stress and the burnout syndrome (p≤0.001; β=0.790). As regards the study hours, they suggest a negative relation with the burnout syndrome (p≤0.05; β=–0.511) and resilience (β=–0.064). By contrast, a positive relation was obtained between the study hours and stress (p≤0.001; β=0.467). Finally, the results point to a negative effect of the burnout syndrome on the levels of resilience (p≤0.001; β=–0.627).

The model proposed for elementary school teacher candidates preparing for the state examination demonstrated good measures of fit for each of its indexes. The chi-square test indicated a non-significant p value (χ2=4.569; df=16; pl=0.001). Similarly, the values of the Comparative Fit Index (CFI), the Goodness of Fit Index (GFI) and the Incremental Fit Index (IFI) were higher than 0.910 while the RMSEA showed a value of 0.079.

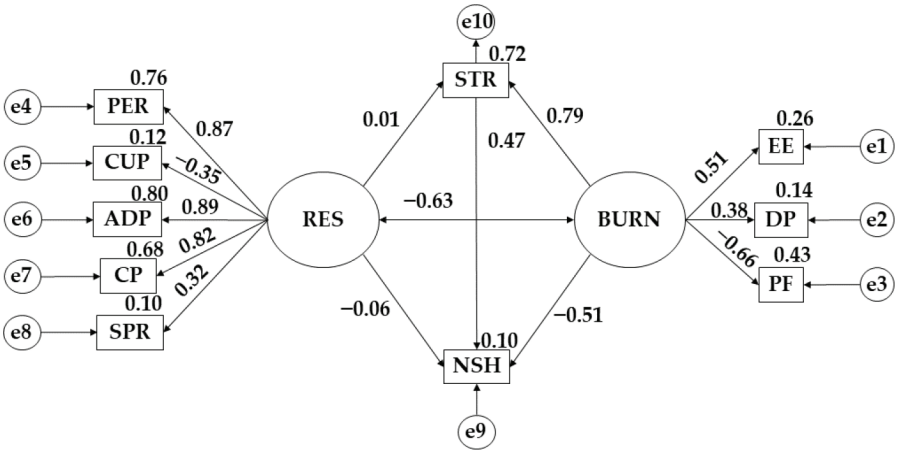
TABLE III. Proposed structural model for preschool teacher candidates

Associations between variables	**R.W.**				**S.R.W.**
	Estimates	**S.E.**	**C.R.**	**Estimates**	**S.E.**
STR ← RES	0.265	2.205	0.120	0.904	0.009
STR ← BURN	1.546	0.223	6.919	***	0.790
PER ← RES	1.000				0.318
CUP ← RES	2.207	0.291	7.585	***	0.824
ADP ← RES	2.180	0.284	7.664	***	0.893
CP ← RES	−1.047	0.181	−5.768	***	−0.346
SPR ← RES	2.471	0.323	7.644	***	0.872
EE ← BURN	1.000				0.506
DP ← BURN	0.680	0.096	7.108	***	0.375
PF ← BURN	−1.315	0.130	−10.135	***	−0.659
NHS ← RES	−0.530	0.647	−0.818	0.413	−0.064
NHS ←BURN	−0.291	0.100	−2.907	**	−0.511
NHS ← STR	0.136	0.035	3.907	***	0.467
RES ← →BURN	−0.615	0.109	−5.657	***	−0.627

Note I: Regression Weights (R.W.); Standardized Regression Weights (S.R.W.); Estimation Error (S.E.); Critical Ratio (C.R.).
Note II: ***$p \leq 0.001$; **$p \leq 0.05$.
Note III: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenaci+ty/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

Received: 19-01-2023 Accepted: 04-03-2023

FIGURE II. Proposed model for preschool teacher candidates



Note: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

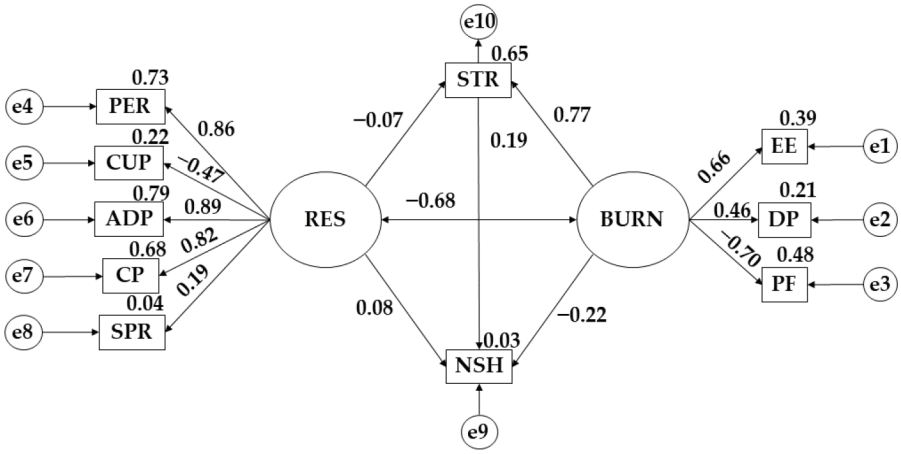
TABLE IV. Results of the proposed model for elementary education candidates

Associations between variables	R.W.				S.R.W.
	Estimates	S.E.	C.R.	Estimates	S.E.
STR ←RES	−3.579	1.562	−2.406	**	−0.072
STR ← BURN	1.347	0.068	19.857	***	0.769
PER ← RES	1.000				0.192
CUP ← RES	3.599	0.349	10.315	***	0.824
ADP ← RES	3.523	0.340	10.351	***	0.890
CP ← RES	−2.447	0.250	-9.784	***	−0.473
SPR ← RES	3.952	0.382	10.335	***	0.857
EE ← BURN	1.000				0.626
DP ← BURN	0.575	0.027	21.438	***	0.455
PF ← BURN	−1.118	0.037	−29.884	***	−0.696
NHS ← RES	1.138	0.448	2.538	**	0.080
NHS ←BURN	−0.106	0.029	−3.659	***	−0.222
NHS ← STR	0.052	0.012	4.265	***	0.190
RES ← →BURN	−0.578	0.061	−9.488	***	−0.681

Note I: Regression Weights (R.W.); Standardized Regression Weights (S.R.W.); Estimation Error (S.E.); Critical Ratio (C.R.).
Note II: ***$p \leq 0.001$; **$p \leq 0.05$.
Note III: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

Received: 19-01-2023 Accepted: 04-03-2023

FIGURE III. Proposed structural model for elementary school teacher candidates



Note: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP) Spirituality (SPR).

The model developed for elementary school teacher candidates highlights the negative effect of resilience on stress ($p \leq 0.05$; β=−0.072). y contrast, a positive effect of the burnout syndrome on stress was observed ($p \leq 0.001$; β=0.769). As regards the daily study hours, the evidence suggests a negative effect of the burnout syndrome ($p \leq 0.001$; β=−0.222) and positive effects of stress ($p \leq 0.001$; β=0.190) and resilience ($p \leq 0.001$; β=0.080). Finally, the results point to a negative relation between resilience and the burnout syndrome ($p \leq 0.001$; β=−0.681).

The model proposed for secondary school teacher candidates preparing for the state examination demonstrated good measures of fit for each of its indexes. The chi-square test indicated a non-significant p value ($\chi 2=5.538$; df=13; pl=0.002). imilarly, the index values were higher than 0.900 while the RMSEA showed a value of 0.088.

The model developed for secondary school teacher candidates evidences the positive effect of resilience and the burnout syndrome on stress (β=0.419; β=0.251). As regards the daily study hours, the evidence suggests a negative effect of the burnout syndrome (β=−0.853) and resilience in like manner (β=−0.481). By contrast, a positive effect of stress on the hours of study was observed (β=0.193). Finally, the results

Received: 19-01-2023 Accepted: 04-03-2023

show a negative effect of resilience on the burnout syndrome ($p \leq 0{,}001$; β=–0.936).

Discussion

Following the presentation of the results above, the discussion that ensues aims to compare them with those of similar studies.

The present descriptive study highlights that secondary school teacher candidates preparing for the state examination have lower levels of stress than pre-school and elementary school teacher candidates. The results of secondary school teacher candidates stand in contrast with

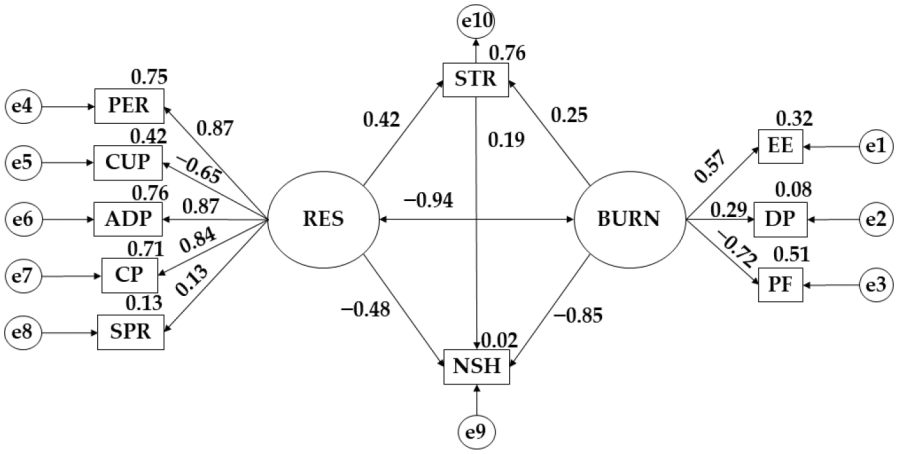
TABLE V. Results of the proposed model for secondary school candidates

Associations between variables	R.W.				S.R.W.
	Estimates	S.E.	C.R.	Estimates	S.E.
STR ←RES	13.860	22.459	0.617	0.537	0.419
STR ← BURN	2.377	1.365	1.741	0.082	0.251
PER ← RES	1.000				0.365
CUP ← RES	2.174	0.307	7.070	***	0.840
ADP ← RES	1.899	0.267	7.120	***	0.872
CP ← RES	−2.182	0.329	−6.635	***	−0.650
SPR ← RES	2.258	0.317	7.113	***	0.867
EE ← BURN	1.000				0.568
DP ← BURN	0.393	0.074	5.299	***	0.289
PF ← BURN	−1.161	0.106	−10.962	***	−0.717
NHS ←BURN	1.000	1.735	−1.012	0.311	−0.853
NHS ← RES	−19.666	20.170	−0.950	0.342	−0.481
NHS ← STR	0.357	0.324	1.102	0.270	0.193
RES ← →BURN	−1.376	0.241	−5.723	***	−0.936

Note I: Regression Weights (R.W.); Standardized Regression Weights (S.R.W.); Estimation Error (S.E.); Critical Ratio (C.R.).
Note II: ***$p \leq 0.001$; **$p \leq 0.05$.
Note III: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

Received: 19-01-2023 Accepted: 04-03-2023

FIGURE IV. Proposed model for secondary school candidates



Note: Stress (STR); Emotional Exhaustion (EE); Depersonalization (DP); Personal Realization (PR); Control and Purpose (CP); Persistence/Tenacity/Self-Efficacy (PER); Control Under Pressure (CUP); Adaptability and Support Networks (ADP); Spirituality (SPR).

those of pre-school and elementary teacher candidates, in that the former group must study a greater number of topics (Orden EDU/3138/2011). Nonetheless, these results may be accounted for the fact that when the data was collected, the state examination for preschool and elementary school teacher candidates was imminent (Resolución de 25 de marzo de 2022). On this matter, the study undertaken by Zandi et al. (2021) indicates that when an examination is approaching, the levels of stress and anxiety increase due to self-imposed expectations to get good grades.

With reference to the burnout syndrome, pre-school and elementary school teacher candidates had higher levels of both mental exhaustion and depersonalization. By contrast, secondary school teacher candidates were found to have a greater sense of personal accomplishment. Concerning these findings, Khoshhal et al. (2017) state that during the preparation for a test, the onset of disruptive states is promoted driving individuals to emotional exhaustion (Ji et al., 2022). Similarly, getting good grades have been found to lead to an increase in personal accomplishment as well as in the academic self-concept (Paechter et al., 2022).

In terms of resilience, secondary school teacher candidates showed higher levels. In relation to these results, Liu et al. (2021) state that when an exam is approaching with long-term preparation, resilience levels are higher. This is in line with research by Yuan (2017), where he states that when exams are approaching, they negatively affect candidates' performance. In addition, candidates in early childhood and primary education were found to study fewer hours. Long-term preparation for an exam leads to increased mental and emotional exhaustion, which has an impact on study hours and, in turn, affects exam preparation in terms of attitude and performance (Melguizo-Ibáñez et al., 2022).

With respect to the effects suggested through the structural equation modeling, the evidence points to a positive effect between resilience and stress among preschool and secondary school teacher candidates. Conversely, a negative effect was observed between both variables among elementary school teacher candidates. Concerning these findings, Anyan et al. (2021) conclude in their study that resilience helps to prevent the effects generated by disruptive states such as anxiety and stress. This effect helps young people to enhance their academic performance and achieve better outcomes (Trigueros et al., 2020). Furthermore, in their study, Ursu and Mairean (2022) claim when a positive relation between disruptive states and resilience is evident, further development in emotion regulation on the subjects' side is needed given that disruptive states exert a stronger effect on resilience.

The results of the present study coincide with those obtained by González-Valero et al. (2022), who assert that the burnout syndrome increases the levels of stress because of the academic and job dissatisfaction that this state generates. Likewise, it has been observed that the burnout syndrome promotes other adverse states for peoples' mental health as it increases the levels of anxiety (Daumiller et al., 2021). The results of the present study also point to a negative effect of the burnout syndrome on the number of study hours among preschool and secondary school teacher candidates. By contrast, they suggest a positive effect among elementary school teacher candidates. Concerning these findings, a study conducted by Olson et al. (2015) concludes that resilience is a factor that benefits the preparation for a test. This capacity helps to mitigate the disruptive states that stem from the preparation process, and thus it helps to enhance performance (Olson et al., 2015).

With regard to the effect of the burnout syndrome and the number of study hours, this study identified a negative effect. Similar results were obtained by Melguizo-Ibáñez et al. (2022), who contend that the burnout syndrome affects the preparation for a test. Similarly, Zheng et al. (2022) conclude that the burnout syndrome promotes emotional exhaustion. A positive effect on the number of study hours was also observed. Results elsewhere found by Agyapong et al. (2022) suggest that undue stress can contribute to resignation from a given activity due to a feeling of distress or incompetence. Finally, a negative relation between the burnout syndrome and resilience was observed in this study. Similar results were obtained by Zhang et al. (2020), who claim that resilience is a mitigating factor that helps to prevent the onset of disruptive states.

Even though this study examines the proposed hypotheses and objectives, it faced a series of limitations that had implications for the investigation.

Firstly, these limitations have to do with the type of study. Given that this study is not longitudinal but cross-sectional, the variables can only be examined at a concrete point in time. The second limitation has to do with the instruments used for the data collection. In spite of the use of validated instruments by the scientific community, they inherently have errors of measurement.

For future lines of research, a program may set out to train teacher candidates sitting the state exam to mitigate disruptive states that stem from the preparation.

Conclusions

The present study investigated the levels of stress, burnout syndrome, resilience of number of study hours in a sample of teacher candidates qualifying for a permanent placement in the public Spanish education system.

The descriptive analysis highlights that pre-school and elementary school teacher candidates preparing for the state examination have higher levels of stress and burnout syndrome than secondary school teacher candidates. By contrast, secondary school teacher candidates show higher levels of resilience and study for more hours.

In respect of the models of structural equations, the results indicate a positive effect of resilience on stress among preschool and secondary

school teacher candidates. By contrast, they show a positive effect of resilience on the number of hours of study among elementary education teacher candidates. Furthermore, they reflect a positive effect of the burnout syndrome on stress. As regards the burnout syndrome, the results point to a negative effect on the number of study hours among all the teacher candidates. With respect to the effect of stress on the number of study hours, a positive relation was observed for all the teacher candidates. Finally, a negative effect of the burnout syndrome on resilience was observed in all the target population of the study.

To conclude, it can be asserted that the training to obtain a permanent teaching post in Spain raises the levels of burnout syndrome and stress among teacher candidates. In light of the findings, the key role of resilience must be highlighted as it prevents the onset of these states and enhances the academic performance in the state exam.

Bibliographic References

Agyapong, B., Obuobi-Donkor, G., Burback, L., & Wei, Y.F. (2022). Stress, Burnout, Anxiety and Depression among Teachers: A Scoping Review. *International Journal of Environmental Research and Public Health, 19*(17), 10706. https://doi.org/10.3390/ijerph191710706

Anyan, F., Ingvaldsen, S.H., & Hjemdal, O. (2021). Interpersonal stress, anxiety and depressive symptoms: Results from a moderated mediation analysis with resilience. *Ansiedad y Estrés, 26*(2-3), 148-154. https://doi.org/10.1016/j.anyes.2020.07.003

Bartell, T., Cho, C., Drake, C., Petchauer, E., & Richmond, G. (2019). Teacher Agency and Resilience in the Age of Neoliberalism. *Journal of Teacher Education, 70*(4), 302-305. https://doi.org/10.1177/0022487119865216

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Calderón, A., Merono, L., & MacPhail, A. (2020). A student-centred digital technology approach: The relationship between intrinsic motivation, learning climate and academic achievement of physical education preservice teachers. *European Physical Education Review, 26*(1), 241-262. http://dx.doi.org/10.1177/1356336X19850852

Received: 19-01-2023 Accepted: 04-03-2023

Cohem, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385-396.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Connor, K.M., & Davidson, J.R.T. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18*(2), 76-82. https://doi.org/10.1002/da.10113

Crespo, M., Fernández-Lansac, V., & Soberón, C. (2014). Spanish version of the Connor-Davidson Resilience Scale (CD-RISC) for chronic stress situations. *Behavioral Psychology, 22*(2), 219-238.

Daumiller, M., Rinas, R., Hein, J., Janke, S., Dickhauser, O., & Dresel, M. (2021). Shifting from face-to-face to online teaching during COVID-19: The role of university faculty achievement goals for attitudes towards this sudden change, and their relevance for burnout/engagement and student evaluations of teaching quality. *Computers in Human Behavior, 118*, 106677. https://doi.org/10.1016/j.chb.2020.106677

De la Fuente, J., & Amate, J. (2019). La experiencia desagradable como determinante de las respuestas cognitivas, conductuales y fisiológicas de estrés académico en universitarios opositores. *Anales de Psicología, 35*(3), 472- 483. https://doi.org/10.6018/analesps.35.3.323101

Díaz-Sánchez, C., & Barra-Almaiga, E. (2017). Resiliencia y satisfacción laboral en profesores de colegios municipales y particulares subvencionados de la comuna de Machalí. *Estudios Pedagógicos, 43*(1), 75-86. http://dx.doi.org/10.4067/S0718-07052017000100005

Freudenberger, H. J. (1989). Burnout: Past, present, and future concerns. *Loss, Grief & Care, 3*(1-2), 1-10.

González-Valero, G., Gómez-Carmona, C.D., Bastida-Castillo, A., Corral-Pernía, J.A., Zurita-Ortega, F., & Melguizo-Ibáñez, E. (2022). Could the complying with WHO physical activity recommendations improve stress, burnout syndrome, and resilience? A cross-sectional study with physical education teachers. *Sport Sciences for Health*, 1-10. https://doi.org/10.1007/s11332-022-00981-6

González-Valero, G., Zurita-Ortega, F., San Román-Mata, & Puertas-Molero, P. (2021). Relación de efecto del Síndrome de Burnout y resiliencia con factores implícitos en la profesión docente. Una revisión sistemática. *Revista De Educación,* (394), 271-296. https://doi.org/10.4438/1988-592X-RE-2021-394-508

Gutentag, T., Horenczyk, G., & Tatar, M. (2017). Teachers' Approaches Toward Cultural Diversity Predict Diversity-Related Burnout and Self-Efficacy. *Journal of Teacher Education, 69*(4), 408-419. https://doi.org/10.1177/0022487117714244

Ji, Y.L., Oubibi, M., Chen, S.Y., Yin, Y.X., & Zhou, Y.L. (2022). Pre-service teachers' emotional experience: Characteristics, dynamics and sources amid the teaching prácticum. *Frontiers in Psychology, 13*, 968513. https://doi.org/10.3389/fpsyg.2022.968513

Khoshhal, K.I., Khairy, G.A., Guraya, S.Y., & Guraya, S.S. (2017). Exam anxiety in the undergraduate medical students of Taibah University. *Medical Teacher, 39*(22). https://doi.org/10.1080/0142159X.2016.1254749

Liu, Y.G., Pan, H.Y., Yang, R.H., Wang, X.J., Rao, J.W., Zhang, X.S., & Pan, C.C. (2021). The relationship between test anxiety and emotion regulation: the mediating effect of psychological resilience. *Annals of General Psychiatry, 20*(1), 40. https://doi.org/10.1186/s12991-021-00360-4

Mansfield, C.F., Beltman, S., Broadley, T., & Weatherby-Fell, N. (2016). Building resilience in teacher education: An evidenced informed framework. *Teaching and Teacher Education, 54*, 77-87. https://doi.org/10.1016/j.tate.2015.11.016

Maslach, C., & Jackson, S.E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior, 2*(2), 99-13. https://doi.org/10.1002/job.4030020205

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin, 107*(2), 247–255. https://doi.org/10.1037/0033-2909.107.2.247

Melguizo-Ibáñez, E., González-Valero, G., Ubago-Jiménez, J.L., & Puertas-Molero, P. (2022). Resilience, Stress, and Burnout Syndrome According to Study Hours in Spanish Public Education School Teacher Applicants: An Explanatory Model as a Function of Weekly Physical Activity Practice Time. *Behavioral Sciences, 12*(9), 329. https://doi.org/10.3390/bs12090329

Olson, K., Kemper, K.J., & Mahan, J.D. (2015). What Factors Promote Resilience and Protect Against Burnout in First-Year Pediatric and Medicine-Pediatric Residents? *Journal of Evidence-Based Integrative Medicine, 20*(3), 192-198. https://doi.org/10.1177/2156587214568894

Orden EDU/3138/2011, de 15 de noviembre, por la que se aprueban los temarios que han de regir en los procedimientos de ingreso, accesos y adquisición de nuevas especialidades de los Cuerpos de Profesores de

Enseñanza Secundaria y Profesores Técnicos de Formación Profesional. *Boletín Oficial del Estado, 278*, de 18 de noviembre de 2011. https://www.boe.es/boe/dias/2011/11/18/pdfs/BOE-A-2011-18099.pdf

Paechter, M., Phan-Lesti, H., Ertl, B., Macher, D., Malkoc, S., & Papousek, I. (2022). Learning in Adverse Circumstances: Impaired by Learning With Anxiety, Maladaptive Cognitions, and Emotions, but Supported by Self-Concept and Motivation. *Frontiers in Psychology, 13*, 850578. https://doi.org/10.3389/fpsyg.2022.850578

Parker, L., Xu, S., & Chi, C. (2022). Chinese Preservice Teachers' Perspectives of Mentoring Relationships in an International Learning Partnership. *Journal of Teacher Education, 73*(5), 525-537. https://doi.org/10.1177/00224871221108655

Raducu, C.M., & Stanculescu, E. (2022). Teachers' Burnout Risk During the COVID-19 Pandemic: Relationships With Socio-Contextual Stress-A Latent Profile Analysis. *Frontiers in Psychiatry, 13*, 870098. https://doi.org/10.3389/fpsyt.2022.870098

Real Decreto 270/2022, de 12 de abril, por el que se modifica el Reglamento de ingreso, accesos y adquisición de nuevas especialidades en los cuerpos docentes a que se refiere la Ley Orgánica 2/2006, de 3 de mayo, de Educación, y se regula el régimen transitorio de ingreso a que se refiere la disposición transitoria decimoséptima de la citada ley, aprobado por Real Decreto 276/2007, de 23 de febrero. *Boletín Oficial del Estado, 88*, de 13 de abril de 2022. https://www.boe.es/eli/es/rd/2022/04/12/270

Remor, E. (2006). Psychometric Properties of a European Spanish Version of the Perceived Stress Scale (PSS). *The Spanish Journal of Psychology, 9*(1), 86-93. https://doi.org/10.1017/S1138741600006004

Resolución de 25 de marzo de 2022, de la Subsecretaría, por la que se convoca procedimiento selectivo de ingreso, por el sistema general de acceso libre y reserva por discapacidad al Cuerpo de Maestros para plazas del ámbito de gestión territorial del Ministerio de Educación y Formación Profesional. *Boletín Oficial del Estado, 76*, de 30 de marzo de 2022. https://www.boe.es/boe/dias/2022/03/30/pdfs/BOE-A-2022-5032.pdf

Seisdedos, N. (1997). *MBI Inventario Burnout de Maslach: Manual*. TEA

Selye, H. (1975). Stress and distress. *Comprehensive Therapy, 1*(8), 9-13.

Suárez-Riveiro, J.M., Rubio-Sánchez, V., Antúnez-Horcajo, R., & Fernández-Suárez, A.P. (2013). Metas y Compromiso de los opositores al

cuerpo de maestros en la especialidad de Educación Primaria. *Revista de Investigación Educativa, 31*(1), 77-92. http://dx.doi.org/10.6018/rie.31.1.139661

Tenenbaum, G., & Eklund, R. (2007). *Handbook of Sport Psychology*. Wiley & Sons.

Tikkanen, L., Haverinen, K., Pyhalto, K., Pietarinen, J., & Soini, T. (2022). Differences in Teacher Burnout Between Schools: Exploring the Effect of Proactive Strategies on Burnout Trajectories. *Frontiers in Education, 7*, 858896. https://doi.org/10.3389/feduc.2022.858896

Trigueros, R., Padilla, A.M., Aguilar-Parra, J.M., Rocamora, P., Morales-Gázquez, M.J., & López-Liria, R. (2020). The Influence of Emotional Intelligence on Resilience, Test Anxiety, Academic Stress and the Mediterranean Diet. A Study with University Students. *International Journal of Environmental Research and Public Health, 17*(6), 2071. https://doi.org/10.3390/ijerph17062071

Ursu, A., & Mairean, C. (2022). Cognitive Emotion Regulation Strategies as Mediators between Resilience and Stress during COVID-19 Pandemic. *International Journal of Environmental Research and Public Health, 19*(19), 12631. https://doi.org/10.3390/ijerph191912631

Yu, T., Li, J.Y., He, L.D., & Pan, X.F. (2022). How Work Stress Impacts Emotional Outcomes of Chinese College Teachers: The Moderated Mediating Effect of Stress Mindset and Resilience. *International Journal of Environmental Research and Public Health, 19*(17), 10932. https://doi.org/10.3390/ijerph191710932

Yuan, R. (2017). Exploring pre-service school counselling teachers' learning needs: perceptions of teacher educators andstudent-teachers. *Journal of Education for Teaching, 43*(4), 474-490. https://doi.org/10.1080/02607476.2017.1342049

Zandi, H., Amirinejhad, A., Azizifar, A., Aibod, S., Veisani, Y., & Mohamadian, F. (2021). The effectiveness of mindfulness training on coping with stress, exam anxiety, and happiness to promote health. *Journal of Education and Health Promotion, 10*(1), 177. https://doi.org/10.4103/jehp.jehp_616_20

Zhang, M., Bai, Y., & Li, Z.Z. (2020). Effect of Resilience on the Mental Health of Special Education Teachers: Moderating Effect of Teaching Barriers. *Psychology Research and Behavior Management, 13*, 537-544. https://doi.org/10.2147/PRBM.S257842

Zhao, W.G., Liao, X.R., Li, Q.T., Jiang, W.N., & Ding, W. (2022). The Relationship Between Teacher Job Stress and Burnout: A Moderated Mediation Model. *Frontiers in Psychology, 12*, 784243. https://doi.org/10.3389/fpsyg.2021.784243

Zhen, B.H., Yao, B.X., & Zhou, X. (2022). Acute stress disorder and job burnout in primary and secondary school teachers during the COVID-19 pandemic: The moderating effect of sense of control. *Current Psychology*. https://doi.org/10.1007/s12144-022-03134-7

Zheng, S., Liu, H.R., & Yao, M.L. (2022). Linking young teachers/self-efficacy and responsibility with their well-being: the mediating role of teaching emotions. *Current Psychology*. https://doi.org/10.1007/s12144-022-03342-1

Contact address: Gabriel González Valero. Universidad de Granada. Facultad de Ciencias de la Educación. Departamento de Didáctica de la Expresión Musical, Plástica y Corporal. Prof. Vicente Callao - Fte Ciencias Educación, 18011, 18011, Granada, Spain. E-mail: ggvalero@ugr.es

Construct Validity of the Gifted Rating Scales (GRS 2) Parent Form in Spain

Validez de Constructo de la Escala de Detección de alumnos con Altas Capacidades para Padres, (GRS 2), en España

https://doi.org/10.4438/1988-592X-RE-2023-402-595

Marta Tourón
https://orcid.org/0000-0001-5430-4198
Universidad Internacional de la Rioja

Enrique Navarro-Asencio
https://orcid.org/0000-0002-3052-146X
Universidad Complutense de Madrid

Javier Tourón
https://orcid.org/0000-0001-8217-1556
Universidad Internacional de la Rioja

Abstract

The development of measurement instruments in the field of high abilities in Spain is scarce. The deficit in the identification of these students is due, in part, to this lack. Current identification procedures focus on a global approach that recommends using various information sources and instruments, such as detection scales, that go beyond intelligence or aptitude test scores, among others. The importance of co-cognitive variables, usually malleable, has been highlighted by many authors, so that the triangulation of complementary information sources is considered essential. Here, for the first time, the study of the construct validity of the GRS 2 Parents Scale in Spain is addressed, with a sample of 1334 fathers and mothers. An exploratory and confirmatory factorial study, AFE and AFC, was carried out, and the metrics of the variables and their multivariate normality have been taken into account, adapting the analysis accordingly. Weighted least squares estimation methods were used. Eight models have been studied

Received: 14-12-2022 Accepted: 12-02-2023

and, finally, a structure of four first-order factors and two second-order factors is proposed, which explains 58% of the variance of the scores. The fit indices of the model are satisfactory (CFI, .98; TLI, .97; GFI, .98), the convergent validity (AVE, above .54 and Composite Reliability, between .78 and .92) show an acceptable result. The proposed structure improves the original three-factor one. This scale is validated in Spain for the first time and provides a measure for the identification of students with high ability. This study will be completed with the validation of the other two scales for teachers that make up the GRS 2, which is being carried out by the authors.

Keywords: gifted rating scales, high ability, gifted identification, construct validity, confirmatory factor analysis.

Resumen

El desarrollo de instrumentos de medida en el ámbito de las altas capacidades en España es escaso. El déficit en la identificación de estos alumnos se debe, en parte, a esta carencia. Los sistemas actuales de evaluación se centran en un enfoque global que recomienda utilizar fuentes de información e instrumentos diversos, como las escalas de detección, que vayan más allá de las puntuaciones de los tests de inteligencia o aptitudes, entre otros. La importancia de las variables co-cognitivas, de ordinario maleables, ha sido puesta de manifiesto por muchos autores, de modo que la triangulación de fuentes de información complementarias se considera esencial. Aquí se aborda por primera vez el estudio de la validez de constructo de la Escala de Padres de las GRS 2 (*Gifted Rating Scales*) en España, con una muestra de 1334 padres y madres. Se llevó a cabo un estudio factorial exploratorio y confirmatorio, AFE y AFC, y se ha tenido en cuenta la métrica de las variables y su normalidad multivariada adecuando los análisis. Se emplearon métodos de estimación de mínimos cuadrados ponderados. Se han estudiado 8 modelos y, finalmente, se propone una estructura de cuatro factores de primer orden y dos factores de segundo orden que explica el 58% de la varianza. Los índices de ajuste del modelo son satisfactorios (CFI, .98; TLI, .97; GFI, .98), la validez convergente (AVE, por encima de .54) y la Fiabilidad Compuesta, con valores entre .78 y .92 muestra un resultado aceptable. La estructura propuesta mejora la original de tres factores. Esta escala es la primera validada en España y aporta una medida para la identificación de los alumnos de alta capacidad relevante. Este estudio se completará con la validación de las otras dos escalas para profesores que componen el GRS 2, y que se está llevando a cabo por los autores.

Palabras clave: escala de detección, altas capacidades, identificación, validez de constructo, análisis factorial confirmatorio.

Introduction

The concept of intelligence understood as a capacity or aptitude and talent as its application to different domains has undergone continuous evolution, from authors such as Galton (1869) or Terman (1925) who described intelligence as a unique, innate and immutable trait, and put their emphasis on its quantification and measurement through IQ tests, to the explanation of a developmental ability, with new models emerging such as those of Gagné (2015, 2018, 2021), Renzulli and Delcourt (2017), Renzulli and Gaesser (2015), Renzulli (2016), Renzulli and Reis (2018, 2021), Pfeiffer (2015a, 2017a) or Olszewski-Kubilius et al. (2019), Worrell et al. (2019) among others, who also highlight the importance of both the context and the co-cognitive factors and variables (Renzulli, 2021) necessary for the development of talent.

On the other hand, the desirability of identifying psychoeducational needs to facilitate the planning of educational care for each schoolchild, in particular for gifted students, and the developmental conception of potential, make it necessary to resort to sources of information beyond the tests associated with the measurement of intelligence or specific aptitudes. Specifically, the importance of context and the non-cognitive or psychosocial, malleable variables proposed by the most contemporary models (Pfeiffer, 2018), brings to the forefront the need for instruments of a diverse nature that make it possible to assess other dimensions beyond the purely intellectual ones, or some dimensions of these, based on standardized tests alone. Thus, the use of different sources of information from the school or family context becomes a necessity (Pfeiffer and Blei, 2008). However, very few instruments are available for this purpose, particularly in Spain.

In this context, it seems particularly important to provide tools to detect student potential and to contribute to a progressive increase in the number of students identified based on their abilities. Although in recent years in Spain some assessment scales for parents, teachers and students have been provided to the educational community, such as those developed by Rogers (2002) and adapted to Spanish by Tourón (2012), which have been available online since 2019, or the Renzulli Scales (Renzulli et al., 2001) for the assessment of the behavioural characteristics of gifted students, the instruments available in Spanish are scarce or lack adequate validation studies.

Received: 14-12-2022 Accepted: 12-02-2023

Therefore, this study aims to carry out the first validation of the parent version of the Gifted Rating Scales (*Gifted Rating Scales. Parent Form, GRS 2*) in Spanish. This newly developed scale belongs to the second version of the GRS, which includes two other scales for teachers, also currently being validated in Spain by the authors of this paper. The GRS 2, particularly the teacher scales, are the ones with the highest coverage, psychometric quality and with the largest amount of research in English conducted.

The *Gifted Rating Scales* were originally designed in 2003 and initially developed to determine whether a student could benefit from participating in a programme for highly gifted students, used in conjunction with other diagnostic tests, as part of a student's comprehensive assessment (Pfeiffer and Jarosewich, 2003). However, a review of the literature prior to the development of the original scales revealed certain shortcomings and limitations in the technical adequacy and usefulness of the existing scales (Jarosewich et al., 2002), which reinforced the need for a screening tool that could be completed by teachers, that would help them to assess observable behaviours of their students that might have high potential, and that would be easy to use, valid and reliable.

Thus, the original GRS included two forms for teachers (GRS-P for students in Preschool/Kindergarten; and GRS-S for students in Primary to Secondary education). This allowed teachers to assess behaviours that might indicate high ability in students aged 4-13 years.

In 2019 the scales used until then were revised and a new scale was added, the parent form (*GRS 2 Parent Form*) (Shaughnessy, 2022).

The GRS 2 parent form, the subject of this paper, is based on a multidimensional model of high abilities and consists of items that fall within three broad dimensions: cognitive abilities, creative and artistic abilities, and social-emotional skills. The items that make up this last dimension are new. The items belonging to the cognitive abilities (8 items) and creative and artistic abilities (8 items) dimensions were adapted from a subset of the items in the original teacher scales, GRS (Pfeiffer and Jarosewich, 2003). Only items reflecting behaviours or characteristics indicative of high ability and observable by a parent outside an educational setting were selected.

The development of the socio-emotional competence dimension was undertaken with the aim of broadening the assessment of the gifted beyond a traditional lens that focuses primarily on “head strengths” –

Received: 14-12-2022 Accepted: 12-02-2023

which include problem solving, memory, reasoning and creativity – to a more holistic and comprehensive view of the student that includes "heart strengths" such as personal and interpersonal strengths (Pfeiffer, 2001, 2017b). Essentially, the purpose was to incorporate a positive psychology perspective into the GRS 2 parenting scale.

The structure of this scale in Spanish has not been studied so far. Precisely, the central aim of this validation work is to analyse the structure of this new scale for parents and to provide evidence on its validity.

Method

Sample

The collaboration of the participants was sought through associations of parents with gifted children (58), other institutions and educational centres, who received a letter explaining the details of the scale and the characteristics of their collaboration. Participation was also requested through various social networks. In all cases collaboration was voluntary and anonymous. The scale was available for online response between April and October 2022.

The sample consisted of a total of 1334 fathers and mothers. After data filtering, the total number of valid responses was 1109; mothers contributed 977 (88.1%) and fathers 113 (10.2%). The remaining 19 cases did not record this information. 61% of them indicated that they were assessing sons and 39% daughters.

The average age of the children assessed by their parents is 10 years (SD= 3.5 years), with cases ranging from 4 to 18 years old. Among them, 55% were in Primary Education (1st-6th grades), 25% in Compulsory Secondary Education (7th-10th grades), 14% in Preschool Education and the remaining 6% in Baccalaureate (11th-12th grades). Table I shows the distribution of cases in the different courses in order of highest to lowest participation.

Naturally, this sample is not representative of all parents with gifted children; however, it should be noted that for the purpose of the study, which is to provide initial evidence of the validity and structure of the scale, a sample of adequate size and variance is sufficient.

TABLE I. Frequencies and percentages of cases in the various courses

Course	Frequencies	%
6th Grade	111	10.0%
4th Grade	105	9.5%
5th Grade	103	9.3%
1st Grade	102	9.2%
2nd Grade	95	8.6%
3rd Grade	94	8.5%
7th Grade	88	7.9%
8th Grade	79	7.1%
2nd Preschool	65	5.9%
9th Grade	62	5.6%
3rd Preschool	59	5.3%
10th Grade	50	4.5%
1st Preschool	33	3.0%
11th Grade	32	2.9%
12th Grade	31	2.8%

Instrument

The *Gifted Rating Scales Second Edition* (GRS 2) is a revision of the original GRS (Pfeiffer and Jarosewich, 2003), integrating the same key elements as its predecessor, but with some new features, including a parent form, which is the one studied here. This parent version of the GRS 2 provides information about the students and allows for scores based on behaviours observed in different contexts, which facilitates a more holistic view of the students' abilities.

The GRS 2 *Parent Form* is a questionnaire for parents of students aged 4-18 years. The inventory consists of a total of 20 items grouped into three dimensions, which are described below.

- **Cognitive Ability**: refers to the child's or teenager's academic abilities, problem solving, reasoning, memory and ability to learn. It encompasses verbal and non-verbal mental abilities, intellectual

Received: 14-12-2022 Accepted: 12-02-2023

abilities or competence, mental speed and the ability to deal with factual or school-related material (Dai, 2018; Gagné, 1993; Olszweski-Kubilius et al., 2019; Sternberg, 1985).

- **Creative/Artistic Ability**: refers to the child's or teenager's ability to think, act or produce unique, original, novel or innovative thoughts or products (Abdulla and Cramond, 2017; Cropley, 2000; Csikszentmihalyi and Wolfe, 2000; Getzels and Jackson, 1962; Isaksen et al., 1993; Plucker and Runco, 1998; Plucker et al., 2018; Runco, 2014). Creative/artistic ability can be expressed in numerous ways: how a child or teenager solves problems, experiments with new ideas (Abdulla and Cramond, 2017; Cropley, 2000), formulates a solution to a group project or uses his or her imagination. Creative children and teenagers are inventive (Beghetto and Plucker, 2016; Cropley and Urban, 2000), curious (Cropley, 2000; Cropley and Urban, 2000; Csikszentmihalyi and Wolfe, 2000; Isaksen et al., 1993; Plucker and Runco, 1998; Presbury et al.,1997), and inquisitive (Presbury et al., 1997; Sternberg, 1985). In addition, children and adolescents gifted in the arts communicate a personalized expressive statement in their work or performance (Hargreaves, 1996; Haroutounian, 1995, 2002; Porath, 1993) through mediums such as art, acting, singing, music, writing or dancing.

- **Socio-Emotional Skills**: refers to the child's or teenager's ability to get along with others, handle stress, remain calm in difficult situations, and be motivated and enthusiastic (Durlak, et al., 2011). Social-emotional skills are not seen as a type of giftedness, but are a measure of a child's or teenager's social skills or competences and emotional intelligence/resilience (Pfeiffer, 2017b). These skills can be observed in a variety of contexts, such as when working cooperatively in a group, coping with challenging situations or dealing with frustrating circumstances (Neihart, et al., 2016).

The dimensions of the scale and its component items are listed in Table II. Each item is rated on a 6-point Likert-type scale: (1) never, (2) sometimes, (3) somewhat often, (4) quite often, (5) almost always and (6) always, according to how parents rate the frequency with which they observe the behaviour indicated in each item (e.g. “Has an active imagination”).

TABLE II. Scale dimensions and items

Dimension	Items
Cognitive Ability	I2, I3, I4, I13, I14, I17
Creative/Artistic Ability	I7, I9, I10, I11, I12, I18, I20
Socio-Emotional Skills	I1, I5, I6, I8, I15, I16, I19

Procedure

The scales, originally in English, were translated independently by the first and third authors, experts in the field of high abilities. Said translations were compared with each other and no discrepancies were observed. Subsequently, the translation was revised by technical staff of the publisher to whom the rights to the scales currently belong (MHS). The authors' proposal was accepted unchanged.

Once the scale was available in Spanish, it was hosted on an online service (Survey Monkey) in order to make it accessible to respondents. The scale items were randomly arranged to avoid possible biases due to the original order of the scales, where items are presented grouped by dimensions (Bishop, 2008; Tourón et al., 2018).

Along with the previous instructions, other descriptive variables were included in the online response form, such as who answers the scale (parent/guardian/other); and with respect to the person assessed: gender, age, grade, performance, psycho-pedagogical evaluation and tests for which information on the child is available.

Data analysis

In order to test the structure of the scale, an exploratory and a confirmatory factor analysis were carried out to provide evidence of the validity of the scale.

To decide on the type of correlations to be used, the normality of the distribution of the responses to each item was tested using Shapiro Wilk's W statistic and the multivariate normality of the set of items was tested using Mardia's symmetry and kurtosis statistics.

Received: 14-12-2022 Accepted: 12-02-2023

On the other hand, a sample of more than 1000 cases, such as the one used in this study, can be considered sufficient. Gaskin and Happell (2014) recommend that, with approximately 6 items per factor, if factor weights of around .5 are obtained, a factor size of approximately 300 cases may be sufficient. With a smaller representation of items, the size should be higher; the same authors recommend about 1000 cases with models that include four items per factor.

The correlations between items are the fundamental information used in factor analysis. The current recommendation with ordinal items, as is the case here, is to use polychoric correlations (Izquierdo et al., 2014; Lloret-Segura et al., 2014). Although with Likert scales of more than 5 points and symmetrical distributions, Pearson's correlation could be used (Viladrich et al., 2017). The method of parameter estimation must also comply with this condition.

Firstly, with 40% of the sample, the structure of the correlations was studied using exploratory factor analysis (EFA) to obtain information on the optimal number of factors. The technique of parallel analysis was used, which estimates the eigenvalues that the factors would obtain by simulating one-dimensional results and compares them with the real data in a sedimentation plot. Secondly, with the rest of the sample, the confirmatory factor analysis (CFA) was carried out. The adequacy of the inter-item correlation matrix was tested using the KMO statistic and the Barlett test. In the first case, values of .8 or more are considered good and, in the second case, a significant result ($p<.05$). These statistics determine whether the size of the inter-item correlations is sufficient to carry out the factor analysis.

Weighted least squares estimation methods were used. And, in the case of CFA, its robust version (WLSM), one of the most recommended options with ordinal variables was used (Li, 2014 and Xia, 2016).

In the model evaluation phase, the standardized indices were used: χ^2 normalized robust (χ^2 /d.f.), to assess the overall fit, where values between 3 and 5 are considered acceptable; RMSEA, to assess the residual matrix, which is acceptable with values below .08; and the TLI (*Tucker-Lewis Index*), for comparative fit, which is acceptable at .90 and above. Following Hu and Bentler (1999), an acceptable fit in the combination of these indices is sufficient as evidence of validity. In addition, the GFI (*Goodness of Fit Index*) and the standardized root mean square error (SRMR) for the overall fit, as well as the CFI (*Comparative Fit Index*) for

the comparative fit are added to the CFA. Modification rates were also calculated in this case.

The dimensionality of the construct is reinforced with the interpretation of the relationship between latent factors. Evidence of convergent validity is also provided through the analysis of factor weights, using the *Average Variance Extracted*, AVE, resulting from adding the standardized factor loadings (Pi) squared (equivalent to R^2) and dividing by the total number of items of the dimension as indicated in formula (1).

$$AVE = \frac{\sum_{i=1}^{N} P_i^2}{n} \qquad (1)$$

In addition, the internal consistency of the dimensions is estimated by calculating the Composite Reliability (CR), based on the factor loadings and the error variance (ei) with the formula (2),

$$FC = \frac{\left(\sum_{i=1}^{n} P_i\right)^2}{\left(\sum_{i=1}^{n} P_i\right)^2 \left(\sum_{i=1}^{n} e_i\right)} \qquad (2)$$

where the error variance of an item is the result of subtracting its squared factor loading from 1, as indicated in formula (3).

$$e_i = 1 - P_i^2 \qquad (3)$$

AVE values of .5 or more indicate that the factor explains 50% or more of the variability of the responses to its component items. Combining this statistic with composite reliability data above .7, convergent validity can be assumed. Information is also included on the explanatory power of the model based on the total variance explained by the set of factors and, in this case, values above 50% are sufficient. The original measurement model, as noted, includes three dimensions (Cognitive Ability, Creative/Artistic Ability and Socio-Emotional Skills). Nonetheless, 1-dimensional and 4-dimensional confirmatory models and some variations included in Table III have been tested, with a total of eight models being tested. Of these, model 4 correlates the errors of items I5 with I15 and I16 with I19; model 5, in addition to correlating these errors, changes item 17 from cognitive to creative ability factor; model 6 proposes a four-factor

Received: 14-12-2022 Accepted: 12-02-2023

TABLE III. Estimated confirmatory models

Model	Structure	Specification
1	3 Factors	Original model
2	4 Factors	The Socio-Emotional Skills factor is subdivided into two dimensions: *Emotional control* and *Social skills*
3	1 Factor	One unique dimension
4	3 Modified A factors	Original model allowing correlations between errors according to modification indexes
5	3 Modified B factors	Original model with correlation between errors and change of item 17 to the Creative dimension
6	4 Modified A factors	Like Model 2 and change item 17 to the Creative dimension
7	3 Modified B factors	Like Model 5 + 1 2nd Order Factors
8	4 Modified A factors	Like Model 6 + 2 2nd Order Factors

structure, changing the dimension of item 17; model 7 is similar to model 5 but with a second-order factor; and finally, model 8 proposes four first-order factors and two second-order factors.

Data analyses have been carried out with jamovi and various modules developed for it (Jamovi Project, 2022; R Core Team, 2021; Friesen et al., 2019; Gallucci and Jentschke, 2021; Rosseel, 2019; Jorgensen et al, 2019 and Epskamp et al. 2019).

Results

The overall mean of the scale (6-point Likert) is 4.5 points (SD=1.22 points) and the median, on average, is 4.7 points. The lowest rated items are 5, 8 and 15 ("Handles stress well"; "Keeps working even when unsuccessful at first" and "Controls his/her anger", respectively), all three with medians of 3 points. On the opposite side, with medians of 6 points, are items 2, 4 and 13 ("Learns things quickly"; "Has a great memory" and "Is quick to understand things", respectively), which also have the lowest variability, with standard deviations of less than 1.

Received: 14-12-2022 Accepted: 12-02-2023

In addition, the correlation of each item with the rest is positive and with values varying between .4 and .7, with an average of .54. A result that indicates the uniformity of the set of items. Their descriptive statistics are shown in table IV.

TABLE IV. Descriptive statistics of the items

Item	Mean	Median	SD	Min.	Max.	Polyserial R
I1	4.47	5	1.36	1	6	.491
I2	5.35	6	0.85	2	6	.561
I3	4.92	5	1.05	2	6	.546
I4	5.36	6	0.92	1	6	.480
I5	2.97	3	1.33	1	6	.439
I6	4.53	5	1.32	1	6	.554
I7	4.80	5	1.24	1	6	.590
I8	3.56	3	1.49	1	6	.553
I9	4.67	5	1.31	1	6	.655
I10	4.62	5	1.24	1	6	.690
I11	4.75	5	1.16	1	6	.659
I12	4.13	4	1.56	1	6	.562
I13	5.37	6	0.87	2	6	.536
I14	5.04	5	1.24	1	6	.431
I15	3.41	3	1.42	1	6	.447
I16	4.01	4	1.32	1	6	.568
I17	4.86	5	1.30	1	6	.489
I18	4.24	4	1.24	1	6	.633
I19	4.75	5	1.10	1	6	.419
I20	5.12	5	1.10	1	6	.615

The results of an analysis of the normality of the variables (items) with the Shapiro-Wilk W test have yielded statistical values of $p<.001$, in all cases, which leads us to reject the null hypothesis. On the other hand, the multivariate normality results for the assumptions of symmetry being distributed as χ^2 (5754.06, $p< .001$) and kurtosis being normally distributed (42.47, $p<.001$), lead us to reject both assumptions.

Received: 14-12-2022 Accepted: 12-02-2023

Given the non-normality of the distribution of the responses, we decided to use the polychoric correlation matrix to carry out the factorial study, which was conducted in two stages.

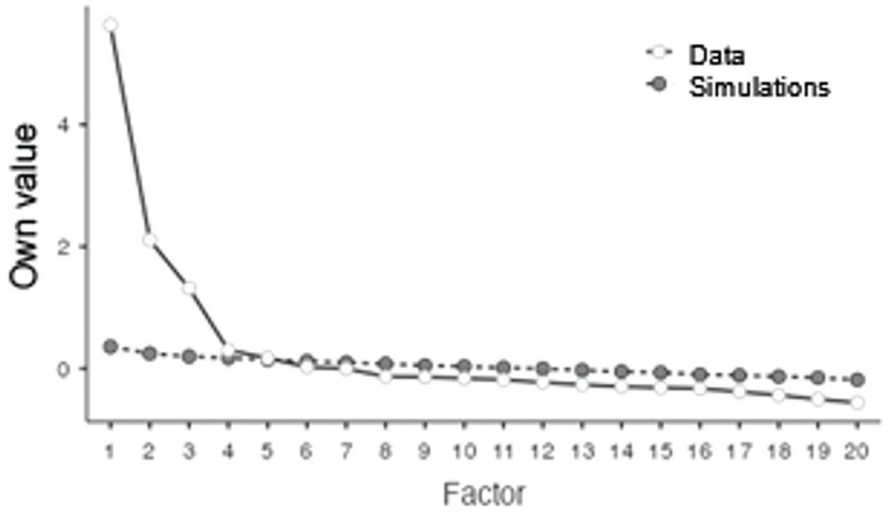
Exploratory Factor Analysis

First, the measures of sampling adequacy (KMO), with an average value of .87 (values above .80 are considered optimal) and a significant Bartlett's test of sphericity result, are indicators that the inter-item correlation values are adequate for the factorial study.

Secondly, the dimensionality results of the AFE indicate a 5-dimensional structure. However, as can be seen in the sedimentation graph of the parallel analysis shown in Figure I, the eigenvalue of the fifth factor is very close to the value of the one-dimensional model resulting from the simulations.

In addition, the factor weights of the fifth factor are from two items (7: "Is creative" and 17: "Gives a lot of detail when explaining things"), which also have a larger load on another factor. On the other hand, the results of the factor weights and the calculated AVE and composite reliability indices show an optimal 4-factor fit (see Table V). Dimension five achieves an AVE of 0.45 and a reliability close to 0.4.

FIGURE I. Sedimentation graph with results of parallel analysis



Received: 14-12-2022 Accepted: 12-02-2023

TABLE V. AVE indices and composite reliability of the four-factor model

Factors	AVE	FC
F1	0.705	0.836
F2	0.703	0.895
F3	0.652	0.750
F4	0.667	0.719

Further evidence in favour of the 4-factor model is its explanatory power. Approximately 60% of the variability in the data is reproduced by it, while the 3-factor model accounts for 54.6%. The fifth factor contributes less than 3% of explained variance. Consequently, the 4-factor model improves the explanation of the items that are located in the fourth factor.

The fit indices of the models estimated in the exploratory study are shown in table VI.

TABLE VI. Exploratory Factor Analysis (EFA) fit indices

Indexes	5 factors	4 factors	3 factors
χ^2	964,885	1347.888	1894.254
gl	100	116	133
p	<.001	<.001	<.001
χ^2/df	9,649	11,620	14,243
RMSEA	.088	.098	.109
RMSEA (LI)	.083	.105	.105
RMSEA (LS)	.093	.114	.114
TLI	.877	.849	.812

Although the results of these indices do not show acceptable values, they seem to improve as the number of factors increases. Considering the evidence of the variability explained by the factors, the factor weights of the items that compose them and the AVE and composite reliability statistics, a 4-dimensional structure would be the one recommended by the EFA.

Received: 14-12-2022 Accepted: 12-02-2023

The items that make up the factors, as a result of the EFA, maintain the original structure, with the only variation being item 17 ("Gives a lot of details when explaining things"), which changes from the Cognitive Ability dimension to the Creative/Artistic Ability one. This is the only difference between the original and the three-factor model.

The four-factor model separates the Socio-Emotional Skills into two sub-dimensions, to place items 5, 8 and 15 in one of them, which, let us recall, are the lowest scoring items of the scale and refer to stress management, anger and perseverance at work. We have named this sub-dimension, *"Emotional Control*". The other sub-dimension groups together items 1, 6, 16 and 19, which, because of their content, we have called *"Social Skills*". These two dimensions, in the original design of the scale, are grouped in the factor that the authors call Socio-Emotional Skills (see table VII).

TABLE VII. Structure of the four-factor EFA model

Factors	Items
Cognitive Ability	I2, I3, I4, I13, I14
Creative/Artistic Ability	I7, I9, I10, I11, I12, I17, I18, I20
Social Skills	I1, I6, I16, I19
Emotional Control	I5, I8, I15

The correlations between the factors, which can be seen in table VIII, show a higher average correlation between Cognitive and Creative/Artistic Abilities on the one hand, and Social and Emotional Skills on the other. The values are not large enough to be considered a single dimension, although they could be explained by second-order common factors. Hypothesis that is tested in the confirmatory stage of the model.

Confirmatory Factor Analysis

Taking into account the results of the EFA, eight confirmatory models were tested in the second stage (see table III), two of them with second-order factors. The fit values for each of the models tested are presented in table IX.

TABLE VIII. Correlations between the factors based on the EFA

	Cognitive Ability	**Creative/Artistic Ability**	***Social Skills***	***Emotional Control***
Cognitive Ability	1			
Creative/Artistic Ability	0.472	1		
Social Skills	0.392	0.385	1	
Emotional Control	0.268	0.136	0.552	1

TABLE IX. Fit indices of models tested in the CFA

Indexes	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
χ^2	2048	1740	6833	1726	1457	1596	1457	1601
gl	167	164	170	165	167	164	165	165
p	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
χ^2/gl	12,263	10,610	40,194	10,461	8,725	9,732	8,830	9,703
SRMR	.086	.078	.154	.075	.071	.074	.075	.074
RMSEA	.101	.093	.189	.089	.084	.089	.084	.089
RMSEA (LI)	.097	.089	.202	.085	.080	.085	.080	.085
RMSEA (LS)	.105	.097	.209	.093	.088	.093	.088	.090
CFI	.968	.974	.867	.976	.978	.976	.978	.976
TLI	.964	.969	.851	.972	.975	.972	.975	.972
GFI	.978	.981	.910	.983	.984	.983	.984	.983

The M1 is the three-factor model (scale's original), where the first factor is Cognitive Ability, the second Creative/Artistic and the third Socio-Emotional Skills. The comparative fit indices (CFI and TLI) achieve good values (>.95). The Goodness of Fit Index (GFI) also points in this direction. In contrast, the residual-based indices (SRMR and RMSEA) slightly exceed acceptable values, exceeding .08. The modification indices suggested an improvement in the fit if the correlations between the residuals of the two groups of items of the Socio-Emotional Skills factor (5, 8 and 15 on the one hand, and 1, 6, 16 and 19 on the other) are included.

Received: 14-12-2022 Accepted: 12-02-2023

This result is further evidence, as already pointed out by the EFA, of the existence of two factors explaining the responses to the Socio-Emotional Skills items.

The M2 tested the grouping of the items into four factors, separating the socio-emotional skills. As can be seen in Table IX, the values of the fit indices improve. Residue rates decrease and there is also a reduction in χ^2 values. And the values of the CFI, TLI and GFI indexes increase.

Subsequently, the M3 tested a unidimensional structure of the construct and, as can be seen, the fit indices worsen considerably.

The M4 and M5 are models with slight variations on the M1. The first includes correlations between the residuals of items 5 and 15 and between 16 and 19, and the second, additionally, shifts item 17 from the cognitive factor to the creative factor. Correlating the residuals assumes the existence of another factor that determines part of the variability of the responses to these items and, as the fit indices show, the results are similar to M2. In addition, changing factor item 17 also produces a slight improvement in the fit values.

The M6 is equivalent to the M2, but by changing item 17 to another dimension and, as with the three-factor models (M5), the fit improves slightly.

Finally, models M7 and M8 include second-order factors considering three and four factors, respectively. M7 links the Cognitive and Creative/Artistic Ability factors and M8 links those two as well as the two Socio-Emotional Skills factors. And, as Table IX shows, the adjustments are roughly equivalent.

Convergent Validity and Reliability

The Convergent Validity (*Average Variance Explained, AVE*) and Composite Reliability presented in Table X show better results in the four-factor models (M2, M6 and M8) with AVE of the Socio-Emotional Skills factors (Factor 3 and Factor 4) with values above .5. In contrast, in the three-factor models (M1, M4, M5 and M7), the Socio-Emotional Skills factor (Factor 3) explains only 40% of the variability.

TABLE X. Evidence of Convergent Validity and Composite Reliability

	M1	M2	M3	M4	M5	M6	M7	M8
Average Variance Explained (AVE)								
Factor 1	.531	.531	.376	.531	.595	.595	.595	.595
Factor 2	.650	.649		.650	.598	.598	.598	.598
Factor 3	.403	.536		.403	.403	.536	.403	.536
Factor 4		.544				.545		.545
Total	.528	.565		.528	.532	.568	.532	.568
Composite Reliability								
Factor 1	.868	.868	.918	.868	.878	.877	.878	.877
Factor 2	.927	.927		.927	.920	.920	.920	.920
Factor 3	.823	.821		.823	.823	.820	.815	.821
Factor 4		.781				.782		.782

The composite reliability values can be considered acceptable in all cases (>.7). The best internal consistency of the items is found in Factor 2 (Creative/Artistic Ability), which achieves reliability values above .9.

The CFA results indicate that the best models, considering the fit values, are M6 and M8. M8 was chosen as the model that best represents the structure of relationships that occur between the items of the construct and allows the scores of the four primary factors and the two second-order dimensions to be used. Table XI includes the factor weights achieved with this model and the proportion of variance explained (R^2).

The factor weights of the items in M8, for the first order factors, are all above .5. The most representative ones are item 2 (b=.873; R^2=.762) in the Cognitive Ability factor, item 10 in Creative/Artistic Ability (b=.919; R^2=.845), item 16 in Social Abilities (b=.838; R^2=.702), and item 15 in Emotional Control (b=.775; R^2=.601). In fact, 70% of the factor weights (b) are above .70 and 30% of them are between .50 and .69.

Received: 14-12-2022 Accepted: 12-02-2023

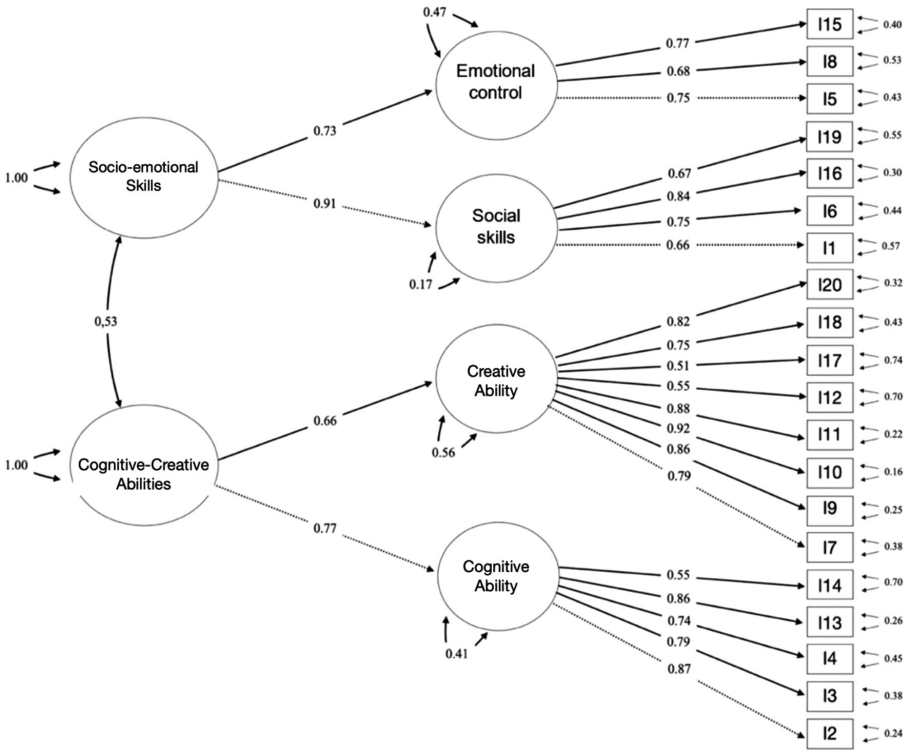
TABLE XI. Factor weights (b), R^2 and residuals of the items and 2nd order factors of Model 8

Dimensions	Items	b	R^2	Error
Cognitive Ability	I2	.873	.762	.238
	I3	.787	.619	.381
	I4	.744	.554	.446
	I13	.861	.741	.259
	I14	.545	.297	.703
Creative/Artistic Ability	I7	.786	.618	.382
	I9	.865	.748	.252
	I10	.919	.845	.155
	I11	.883	.780	.220
	I12	.546	.298	.702
	I17	.505	.255	.745
	I18	.752	.566	.434
	I20	.822	.676	.324
Social Skills	I1	.657	.432	.568
	I6	.751	.564	.436
	I16	.838	.702	.298
	I19	.668	.446	.554
Emotional Control	I5	.752	.566	.434
	I8	.684	.468	.532
	I15	.775	.601	.399
2nd order Factors	Dimensions	b	R^2	Error
Cognitive-Creative Ability	Cognitive	.767	.588	.412
	Creative	.662	.438	.562
Socio-Emotional Skills	*Social Skills*	.913	.834	.166
	Emotional Control	.727	.529	.471

The second-order factors also explain more than 50% of the variability of the scores; specifically, 51.3% in the case of the Cognitive-Creative factor and 68.1% in the Socio-Emotional Skills factor, which reinforces the superiority of this model over the others.

Finally, the plot of the effects between factors and items of M8 is shown in Figure II.

FIGURE II. Diagram of observed, latent variables and effects in the M8



Conclusions

This is the first study carried out on the factor structure of this instrument in Spanish. The only reference for assessing the results obtained is the original structure of the scale, organized in three dimensions as explained above.

After testing several models and considering the above fit and error indices, we have found that a first-order four-factor structure is the most suitable. This structure involves dividing the original Socio-Emotional Skills dimension into two parts, which we have named based on the content of the items included: *Social Skills* and *Emotional Control*. Thus, we would have a four-factor structure instead of a three-factor structure. Further analysis has led us to verify that the fit to the sample data improves

Received: 14-12-2022 Accepted: 12-02-2023

when considering, in addition, two second-order factors that group cognitive and creative artistic abilities, on the one hand, and the other factor that groups the social and emotional control dimensions, which we call Socio-Emotional Skills (Model 8).

From a practical point of view, this division of the Socio-Emotional Skills dimension into two components (*Social Skills* and *Emotional Control*) may refine the assessment of the candidates' profile by distinguishing these two sub-dimensions (factors), although the proposed measurement model would also allow for the integration of the two dimensions in case homogeneous scores are obtained in both and one does not mask the other.

The need and desirability of including parents as a source of information in the process of identifying and assessing gifted students, which can be triangulated with other data sources, has been highlighted by several authors (Nicpon and Pfeiffer, 2011; Pfeiffer, 2015b), and the lack of instruments has been seen as a gap in this field. On the other hand, parents can provide relevant data from contexts other than school, in particular from the socio-emotional domain.

This dimension reflects what Pfeiffer (2001, 2017b) has called the "strengths of the heart", which go beyond cognitive variables. In fact, other authors have highlighted the importance of determination, courage, hard work, etc. as essential dimensions in talent development (Duckworth, 2016; Pfeiffer, 2013; Subotnik et al., 2021; Olszewski-Kubilius et al., 2015, 2019).

Pfeiffer (2015b) notes that some authors argue that parents' assessments may be biased in their perceptions of their children's skills and abilities; however, research indicates that identification systems that include parental nominations improve and prevent large numbers of gifted students from going unidentified (McBee et al., 2016). In fact, there are studies that show the validity of parental assessments and their correlation with measures of aptitude and performance (Lee and Olszewski-Kubilius, 2006), and numerous authors recommend including these assessments in a process of global *screening* that, together with those from teachers, allow us to learn about aspects of students not covered by intelligence and aptitude tests (Pfeiffer, 2017a).

For all these reasons, it is of the utmost importance to have properly validated instruments, such as the one presented here, so that they can be used in educational practice.

Received: 14-12-2022 Accepted: 12-02-2023

This work is certainly a preliminary validation that will require further confirmatory studies, both with Spanish samples and comparative studies with other similar investigations carried out in different cultural contexts.

The official figures for students identified as having high intellectual ability in Spain are quite concerning, as they indicate that nearly 98% of students go undetected (Tourón, 2020). In light of this, the availability of adequately validated instruments is particularly relevant.

This work will be completed with the validation of the other two teacher scales that make up GRS 2, which is being carried out by the authors. This will represent a significant advance in the use of useful tools for the identification of students' potential.

Funding

This research was supported by the Research Vice-rectorate of Universidad Internacional de La Rioja.

Bibliographic References

Abdulla, A. M., & Cramond, B. (2017). After Six Decades of Systematic Study of Creativity: What do Teachers Need to Know About What it is and How it is Measured? *Roeper Review, 39*(1), 9-23.

Beghetto, R. A., & Plucker, J. A. (2016). Revisiting the Relationship Among Schooling, Learning, and Creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity and Reason in Cognitive Development* (pp. 72–91). Cambridge University Press. https://doi.org/10.1017/CBO9781139941969.005

Bishop, G. F. (2008). Item Order Randomization. In P. J. Lavrakas, *Encyclopedia of Survey Research Methods* (p. 397). SAGE.

Cropley, A. J. (2000). Defining and Measuring Creativity: Are Creativity Tests Worth Using? *Roeper Review, 23*(2), 72-79.

Cropley, A. J., & Urban, K. K. (2000). Programs and Strategies for Nurturing Creativity. In K. A. Heller, F. J. Mönks, R. Subotnik, Robert Sternberg (Eds.), *International Handbook of Giftedness and Talent* (2nd ed., pp. 485-498). Pergamon Press.

Csikszentmihalyi, M., & Wolfe, R. (2000). New Conceptions and Research Approaches to Creativity: Implications of a Systems Perspective for Creativity in Education. In K. A. Heller, F. J. Mönks, R. Subotnik, Robert Sternberg (Eds.), *International Handbook of Giftedness and Talent* (2nd ed., pp. 81-93). Pergamon Press.

Dai, D. Y. (2018). A History of Giftedness: Paradigms and Paradoxes. In S. I. Pfeiffer (Ed.), *Handbook of Giftedness in Children* (2nd. ed., pp. 1-14). Springer. https://doi.org/10.1007/978-3-319-77004-8_1

Duckworth, A. (2016). *Grit: The Power of Passion and Perseverance*. Scribner/Simon & Schuster.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development, 82*(1), 405-432.

Epskamp S., Stuber S., Nak J., Veenman M, Jorgensen T. D. (2019). *semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output*. [R Package]. Retrieved from https://CRAN.R-project.org/package=semPlot

Friesen, L., Kroc, E., Zumbo, B. D. (2019). *Psychometrics & Post-Data Analysis: Non-Parametric Item Response Theory*. [jamovi module]. Retrieved from https://github.com/lucasjfriesen/jamoviPsychoPDA

Gagné, F. (1993). Constructs and Models Pertaining to Exceptional Human Abilities. In K. A. Heller, F. J. Mönks, & A. H. Passow (Eds.), *International Handbook of Research and Development of Giftedness and Talent* (pp. 69–87). Pergamon Press.

Gagné, F. (2015). From Genes to Talent: The DMGT/CMTD Perspective. *Revista de Educación, 368*, 66-91. https://doi.org/10.4438/1988-592X-RE-2015-368-289

Gagné, F. (2018). Academic Talent Development: Theory and Best Practices. In S. I. Pfeiffer, E. Shaunessy-Dedrick, & M. Foley-Nicpon (Eds.), *APA Handbook of Giftedness and Talent* (pp. 163–183). American Psychological Association. https://doi.org/10.1037/0000038-011

Gagné, F. (2021). The Differentiating Model of Giftedness and Talent. In T. L. Cross & J. R. Cross (Eds.), *Handbook for Counselors Serving Students with Gifts and Talents: Development, Relationships, School Issues, and Counseling Needs/Interventions* (2nd ed., pp. 9-28). Routledge. https://doi.org/10.4324/9781003235415-3

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

Galton, F. (1869). *Hereditary Genius: An Inquiry Into its Laws and Consequences*. Macmillan and Co.

Gaskin, C. J., & Happell, B. (2014). On Exploratory Factor Analysis: A Review of Recent Evidence and Assessment of Current Practice, and Recommendations for Future Use. *International Journal of Nursing Studies, 51*(3), 511-521.

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and Intelligence: Explorations with Gifted Students*. Wiley.

Hargreaves, A. (1996). Revisiting Voice. *Educational Researcher, 25*(1), 12-19.

Haroutounian, J. (1995). Talent Identification and Development in the Arts: An Artistic/Educational Dialogue. *Roeper Review, 18*(2), 112-117.

Haroutounian, J. (2002). *Kindling the Spark: Recognizing and Developing Musical Talent*. Oxford University Press on Demand.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria *versus* New Alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Isaksen, S. G., Puccio, G. J., & Treffinger, D. J. (1993). An Ecological Approach to Creativity Research: Profiling for Creative Problem Solving. *The Journal of Creative Behavior, 27*(3), 149–170. https://doi.org/10.1002/j.2162-6057.1993.tb00704.x

Izquierdo, I., Olea, J., & Abad, F. J. (2014). El análisis factorial exploratorio en estudios de validación: Usos y recomendaciones [Exploratory Factor Analysis in Validation Studies: Uses and Recommendations]. *Psicothema, 26*(3), 395-400.

Jamovi Project, The (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from https://www.jamovi.org

Jarosewich, T., Pfeiffer, S. I., & Morris, J. (2002). Identifying Gifted Students Using Teacher Rating Scales: A Review of Existing Instruments. *Journal of Psychoeducational Assessment, 20*(4), 322-336.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., Byrnes, J., Rights, J. D., Longo, Y., Mansolf, M., Ben-Shachar, M. S., Rönkkö, M. (2019). *semTools: Useful Tools for Structural Equation Modeling*. [R Package]. Retrieved from https://CRAN.R-project.org/package=semTools

Received: 14-12-2022 Accepted: 12-02-2023

Lee, S. Y. & Olszewski-Kubilius, P. (2006). Comparison Between Talent Search Students Qualifying via Scores on Standardized Tests and via Parent Nomination. *Roeper Review, 28*, 157-166.

Li, Ch.-H. (2014). The Performance of MLR, USLMV, AND WLSMV Estimation in Structural Regression Models with Ordinal Variables. Doctoral Dissertation. Michigan State University.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. [Exploratory Item Factor Analysis: a practical guide revised and updated], *Anales de Psicología/Annals of Psychology, 30*(3), 1151-1169.

McBee, M. T., Peters, S. J., & Miller, E. M. (2016). The Impact of the Nomination Stage on Gifted Program Identification: A Comprehensive Psychometric Analysis. *Gifted Child Quarterly, 60*(4), 258–278. https://doi.org/10.1177/0016986216656256

Neihart, M., Pfeiffer, S., & Cross, T. (2016). *The Social and Emotional Development of Gifted Children: What Do We Know?* (2nd ed.). Prufrock.

Nicpon, M. F., & Pfeiffer, S. I. (2011). High-Ability Students: New Ways to Conceptualize Giftedness and Provide Psychological Services in the Schools. *Journal of Applied School Psychology, 27*(4), 293-305.

Olszewski-Kubilius, P., Subotnik, R. & Worrell, F. C. (2015). Re- pensando las altas capacidades: una aproximación evolutiva. [Re-Thinking Giftedness: A Developmental Approach]. *Revista de Educación, 368*, 40-65. https://doi.org/10.4438/1988-592X-RE-2015-368-297

Olszewski-Kubilius, P., Worrell, F. C., Subotnik, R. F., & Foley-Nicpon, M. (2019). Reflections on Talent Development in Academics. In R. F. Subotnik, P. Olszewski-Kubilius, & F. C. Worrell (Eds.), *The Psychology of High Performance: Developing Human Potential into Domain-Specific Talent* (pp. 225–236). American Psychological Association. https://doi.org/10.1037/0000120-011

Pfeiffer, S. I. (2001). Professional Psychology and the Gifted: Emerging Practice Opportunities. *Professional Psychology: Research and Practice, 32*(2), 175.

Pfeiffer, S. I. (2013). Lessons Learned from Working with High-Ability Students. *Gifted Education International, 29*(1), 86-97.

Pfeiffer, S. I. (2015a). Tripartite Model of Giftedness and Best Practices in Gifted Assessment. *Revista de Educación, 368,* 66-95. DOI: 10.4438/1988-592X-RE-2015-368-293

Pfeiffer, S. I. (2015b). *Essentials of Gifted Assessment*. John Wiley & Sons.

Pfeiffer, S. I. (2017a). *Identificación y evaluación del alumnado con altas capacidades: Una guía práctica*. [Identification and Evaluation of High Ability Students: A Practical Guide]. UNIR Editorial.

Pfeiffer, S. I. (2017b). Success in the Classroom and in Life: Focusing on Strengths of the Head and Strengths of the Heart. *Gifted Education International, 33*(2), 95-101.

Pfeiffer, S. I. (2018). Giftedness and Talent Development in Children and Youth. In P. Ward, J. M. Schraagen, J. Gore, & E. M. Roth (Eds.), *The Oxford Handbook of Expertise* (pp. 103-127). https://doi.org/10.1093/oxfordhb/9780198795872.013.5

Pfeiffer, S. I., & Blei, S. (2008). Gifted Identification Beyond the IQ test: Rating Scales and Other Assessment procedures. In S. I. Pfeiffer (Ed.), *Handbook of Giftedness in Children: Psychoeducational Theory, Research, and Best Practices* (pp. 177–198). Springer Science. https://doi.org/10.1007/978-0-387-74401-8_10

Pfeiffer, S. I., & Jarosewich, T. (2003). *GRS: Gifted Rating Scales*. Psychological Corporation.

Plucker, J.A., Guo, J., & Makel, M. C. (2018). Creativity. In S. I. Pfeiffer (Ed.), *Handbook of Giftedness in Children* (pp. 81-99). Springer. https://doi.org/10.1007/978-3-319-77004-8_6

Plucker, J. A., & Runco, M. A. (1998). The Death of Creativity Measurement Has Been Greatly Exaggerated: Current Issues, Recent Advances, and Future Directions in Creativity Assessment. *Roeper Review, 21*(1), 36-39.

Porath, M. (1993). Gifted Young Artists: Developmental and Individual Differences. *Roeper Review, 16*(1), 29-33.

Presbury, J. H., Benson, A. J. & Torrance, E. P. (1997). Creativity. In G. G. Bear, K. M. Minke, & A. Thomas (Eds.), *Children's needs II: Development, Problems and Alternatives* (pp. 449-456). National Association of School Psychologists.

R Core Team (2021). *R: A Language and environment for statistical computing*. (Version 4.1) [Computer software]. https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2022-01-01).

Renzulli, J. S. (2016). The Three-Ring Conception of Giftedness. In J. Renzulli, S. M. Reis*, Reflections on Gifted Education: Critical Works by Joseph S. Renzulli and Colleagues* (pp. 55-90). Prufrock. https://doi.org/10.4324/9781003237693-5

Received: 14-12-2022 Accepted: 12-02-2023

Renzulli, J. S. (2021). Intelligences Outside the Normal Curve: Co-Cognitive Factors That Contribute to the Creation of Social Capital and Leadership Skills in Young People. In J. Renzulli, S. M. Reis, *Reflections on Gifted Education: Critical Works by Joseph S. Renzulli and Colleagues* (pp. 129-150). https://doi.org/10.4324/9781003237693-5

Renzulli, J. S., & Delcourt, M. A. B. (2017). Gifted Behaviors *versus* Gifted Individuals. In C. M. Callahan, H. L. Hertberg-Davis, *Fundamentals of Gifted Education* (pp. 42-54). Routledge. https://doi.org/10.4324/9781315639987-5

Renzulli, J. S., & Gaesser, A. H. (2015). A Multi Criteria System for the Identification of High Achieving and Creative/Productive Giftedness. *Revista de Educación, 368*, 96-131. https://doi.org/10.4438/1988-592X-RE-2015-368-290

Renzulli, J. S., Hartment, R., & Callahan, C. (2001). *Escalas de Renzulli (SCRBSS): Escalas para la valoración de las características de comportamiento de los estudiantes superiores* [Scales for Rating the Behavioral Characteristics of Superior Students]. Amarú Ediciones.

Renzulli, J. S., & Reis, S. M. (2018). The Three-Ring Conception of Giftedness: A Developmental Approach for Promoting Creative Productivity in Young People. In S. I. Pfeiffer, E. Shaunessy-Dedrick, & M. Foley-Nicpon (Eds.), *APA Handbook of Giftedness and Talent* (pp. 185–199). American Psychological Association. https://doi.org/10.1037/0000038-012

Renzulli, J. S., & Reis, S. M. (2021). The Three Ring Conception of Giftedness: A Change in Direction from Being Gifted to the Development of Gifted Behaviors. In R. J. Sternberg & Ambrose, D. (Eds.), *Conceptions of Giftedness and Talent* (pp. 335-355). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-56869-6_19

Rogers, K. B. (2002). *Re-forming Gifted Education: Matching the Program to the Child*. Great Potential Press, Inc.

Rosseel, Y. (2019). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48(2), 1-36. DOI:* https://doi.org/10.18637/jss.v048.i02

Runco, M. A. (2014). "Big C, Little c" Creativity as a False Dichotomy: Reality is not Categorical. *Creativity Research Journal, 26*(1), 131-132.

Shaughnessy, M. F. (2022). An Interview with Steven I. Pfeiffer: Gifted Rating Scales-(GRS [TM] 2). *North American Journal of Psychology, 24*(1), 1-1.

Sternberg, R. J. (1985). Implicit Theories of Intelligence, Creativity, and Wisdom. *Journal of Personality and Social Psychology, 49*(3), 607.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2021). Unlocking Creative Productivity: A Talent Development Approach. *Journal of Modern Foreign Psychology, 10*(4). https://doi.org/10.17759/jmfp.2021100402

Terman, L. M. (1925). Mental and Physical Traits of a Thousand Gifted Children. *Genetic Studies of Genius*. Volume 1. Stanford University Press.

Tourón, J. (2012). Los cuestionarios de Rogers para la nominación de los más capaces. Padres y Profesores [Rogers Questionnaires for Parent and Teacher Gifted Nominations]. https://www.javiertouron.es/los-cuestionarios-de-rogers-para-la

Tourón, J. (2020). Las altas capacidades en el sistema educativo español: reflexiones sobre el concepto y la identificación [Gifted Students in the Spanish Education System: Reflections on Conceptualization and Identification Issues]. *Revista de Investigación Educativa, 38*(1), 15-32.

Tourón, J., Martín, D., Navarro Asensio, E., Pradas, S. e Íñigo, V. (2018). Validación de constructo de un instrumento para medir la competencia digital docente de los profesores (CDD) [Construct validation of a questionnaire to measure teachers' digital competence (TDC)]. *Revista Española de Pedagogía, 76*(269), 2554. https://doi.org/10.22550/REP761201802

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A Journey Around Alpha and Omega to Estimate Internal Consistency Reliability. *Annals of Psychology, 33*(3), 755-782. https://doi.org/10.6018/analesps.33.3.268401

Worrell, F. C., Subotnik, R. F., Olszewski-Kubilius, P., & Dixson, D. D. (2019). Gifted Students. *Annual Review of Psychology, 70*, 551-576.

Xia, Y. (2016). Investigating the Chi-square Based Model-fit Indexes for WLSMV and ULSMV Estimators (Dissertation). Florida State University. http://purl.flvc.org/fsu/fd/FSU_2016SU_Xia_fsu_0071E_13379

Contact address: Marta Tourón. Universidad Internacional de La Rioja – UNIR, Avd. de La Paz, 137. 26006 Logroño La Rioja, Spain. E-mail: marta.tporto@unir.net

Received: 14-12-2022 Accepted: 12-02-2023

Training in the Primary Education Teacher Degree in STEM disciplines: Analysis before its reform in Spain

Formación en el Grado de Educación Primaria en las disciplinas STEM: Análisis ante su reforma en España

https://doi.org/10.4438/1988-592X-RE-2023-402-596

Germán Ros Magán
https://orcid.org/0000-0001-6623-1483
Universidad de Alcalá

Iñigo Rodríguez Arteche
https://orcid.org/0000-0001-7904-6682
Universidad de Alcalá

Arántzazu Fraile Rey
https://orcid.org/0000-0002-2897-3029
Universidad de Alcalá

Julio Pastor Mendoza
https://orcid.org/0000-0003-0032-4446
Universidad de Alcalá

Abstract

The new Law of Education in Spain, known as LOMLOE, includes for the first time the STEM competence with its own entity, among other changes. Accordingly, a reform of the curricula of the Primary Education Teacher Degrees is proposed, for which a deep and global analysis of this initial training is essential. Thus, the objectives that guided this study included the following: *i)* to compare the distribution of credits among STEM disciplines and their didactics; *ii)* to check whether scientific-mathematical or didactic content predominates; *iii)* to analyse the contents of Technology and Engineering; and *iv)* to compare these aspects between public and private universities. To that end, the study plans

Received: 02-08-2022 Accepted: 02-04-2023

of the 37 public and 22 private universities that offered the Primary Education Teacher Degree in the 2020-2021 academic year were analysed. The teaching guides of the 342 compulsory subjects of the associated disciplines were evaluated, considering their contents and their timing, and a descriptive and inferential study was carried out. The results show that public universities allocate more credits to training in STEM disciplines, especially to the disciplinary content of mathematics and science, while there is a significant dispersion in the data. In public universities there is a major focus on disciplinary content, especially in science. Regarding technology, the contents are specific to Information and Communication Technologies, and not to technology understood from the STEM perspective. In addition, engineering and the global STEM approach itself are absent. This analysis reveals the need to renew the curricula with contents that are more adapted to what the LOMLOE demands and, above all, explicitly including aspects of technology and engineering with a STEM approach.

Keywords: STEM, LOMLOE, pre-service primary teacher training, science, mathematics, technology, engineering, public universities, private universities.

Resumen

La nueva Ley de Educación, conocida como LOMLOE, incluye entre otros cambios y por vez primera la competencia STEM con entidad propia. En consonancia, se propone una reforma de los planes de estudio de los Grados de Magisterio en Educación Primaria, para lo cual resulta indispensable un análisis profundo y global de dicha formación inicial. En este estudio se abordan los siguientes objetivos: *i)* comparar la distribución de créditos en las disciplinas STEM y sus didácticas; *ii)* comprobar si predominan los contenidos científico-matemáticos o los didácticos; *iii)* analizar los contenidos de Tecnología e Ingeniería; *iv)* comparar estos aspectos entre universidades públicas y privadas. Para ello, se revisan los planes de estudios de las 37 universidades públicas y 22 privadas que ofertan el Grado de Magisterio en Educación Primaria en el curso 2020-2021. Se analizan las guías docentes de las 342 asignaturas obligatorias de las disciplinas relacionadas, atendiendo a sus contenidos y a su temporalización, y se realiza un estudio descriptivo e inferencial. Los resultados muestran que las universidades públicas dedican más créditos a la formación en las disciplinas STEM, especialmente a los contenidos disciplinares de matemáticas y ciencias, existiendo una importante dispersión en los datos. En las universidades públicas, los contenidos disciplinares predominan sobre los didácticos, especialmente en ciencias. En cuanto a la tecnología, los contenidos son propios de las llamadas Tecnologías de la Información y la Comunicación, y no de la tecnología entendida desde STEM. Por su parte, la ingeniería y el propio enfoque global STEM están ausentes. Este análisis revela la necesidad de renovar los planes de estudio con contenidos adaptados a lo que demanda la LOMLOE, fundamentalmente, incluyendo aspectos de tecnología e ingeniería desde una perspectiva STEM.

Received: 02-08-2022 Accepted: 02-04-2023

Palabras clave: STEM, LOMLOE, formación inicial de maestros, ciencias, matemáticas, tecnología, ingeniería, universidades públicas, universidades privadas.

Introduction

The international expansion of the STEM educational approach (acronym for Science, Technology, Engineering, and Mathematics) has been unstoppable since its conception in the United States in the 1990s. Nowadays, its influence is evident in most educational contexts across the globe (Belbase et al., 2021). Although there is no clear conceptualisation of STEM, Moore et al. (2014) define STEM education as "an effort to combine some or all of the four disciplines of science, technology, engineering, and mathematics into one class, unit, or lesson that is based on connections between the subjects and real-world problems" (p.38). On this basis, there are different perspectives on integration in STEM, ranging from those that advocate for the fusion of all disciplines through transdisciplinary problems (Costantino, 2018) to others that emphasise the uniqueness of each subject while acknowledging its epistemic complementarity (Simarro and Couso, 2021). In addition, in recent years, there has also been a demand for the integration of Arts and Humanities into this approach (which could become STEAM), to unify convergent and divergent thinking and promote more inclusive academic and professional identities (Aguilera and Ortiz-Revilla, 2021).

The Spanish prospective report titled "Spain 2050" (Ministry of the Presidency, 2021) highlights the importance of increased training in STEM areas. Therefore, the new Education Law, recently approved and known as LOMLOE, introduces the STEM competence for the first time (mathematical competence and competence in science, technology, and engineering), which "involves understanding the world using scientific methods, mathematical thinking and representation, technology, and engineering methods to transform the environment in a committed, responsible, and sustainable way" (Royal Decree 157/2022, p.24406). Moreover, this key competence is associated with five operational descriptors in both Primary and Secondary Education.

With the introduction of this new legal framework, the Ministry of Education has also proposed updating Primary Education Teacher Degrees (MEFP, 2022a), given the critical role of teachers' practice in promoting scientific literacy. Therefore, if the aim is to improve teacher

Received: 02-08-2022 Accepted: 02-04-2023

training for the development of the STEM competence, a prior analysis of initial training in the various STEM disciplines in Primary Education Teacher Degrees is required. Although there are several studies for each discipline, a joint analysis such as the one proposed in this paper has not been conducted yet.

Furthermore, a proper analysis should include private universities, as they represent a high percentage of students in Spain, specifically 26.6% in the 2020/21 academic year (MEFP, 2022b). Previous studies that have analysed the overall curriculum, i.e., the distribution of credits between basic, disciplinary, and didactic subjects, field experiences, and the final degree project, have not found significant differences with public universities (Sánchez-Urán, 2019). However, each university decides how to distribute the credits of each module in the different subjects and what content to include. This has been criticised, as these decisions are sometimes based on the trajectory of each faculty or the interests of the departments, rather than on the benefit of the students (Imbernón and Colén, 2014). Moreover, 30 out of the total 240 ECTS of the bachelor's degree are designated as electives and are normally devoted to elective courses or majors.

Training in STEM disciplines

According to the "Teacher Education and Development Study in Mathematics" (INEE, 2012), primary preservice teachers lack mathematical knowledge and skills, leading to decreased confidence, motivation, enjoyment, and sense of the subject's usefulness (Nortes Checa and Nortes Martínez-Artero, 2013). These deficiencies are attributed in part to inadequacies in their previous training and, potentially, to an overemphasis on operational thinking (Socas, 2011). Similar conclusions can be drawn from science education. As such, various national (Martínez-Borreguero et al., 2022) and international authors (Appleton, 2003) have found that future primary school teachers often have inadequate scientific knowledge, alternative conceptions similar to those of school-age pupils, and a rigid understanding of science, which hinders exploration and positive attitudes towards the subject (Porlán et al., 2010).

On the other hand, there is a broad consensus regarding the indispensability of adequate Pedagogical Content Knowledge (PCK hereafter) in both Mathematics and Science, since this element is essential to represent

Received: 02-08-2022 Accepted: 02-04-2023

or transform disciplinary knowledge in such a way that students can understand it (Shulman, 1987). In both disciplines, PCK is assumed to comprise knowledge of curricula, students' understanding of these subjects, teaching methods, assessment strategies, as well as orientations towards teaching (Cortés et al., 2012; Naya- Riveiro et al., 2021). In Mathematics, both national and foreign authors (Cardetti and Truxaw, 2014; Naya-Riveiro et al., 2021) highlight the deficiencies in university training regarding mathematical knowledge for teaching. The situation is similarly unsatisfactory in Science, where Cañal (2008) and Toma et al. (2017) analysed the design of inquiry-based teaching units by primary preservice teachers and identified their difficulties in integrating conceptual and procedural knowledge, promoting student reflection, and providing adequate scaffolding.

Although there is widespread recognition of the need to improve the scientific-mathematical and didactic knowledge of future teachers, there is no consensus on whether the didactic component should follow conceptual knowledge per se or whether both should be addressed simultaneously (Amat et al., 2022). In fact, tensions arise between what could be described as a more scientific approach and a more didactic one, and doubts persist on how to bring the two together (García-Barros, 2016). On the one hand, the lack of conceptual knowledge hinders PCK development (Cañal, 2008; Martínez-Borreguero et al., 2022). On the other hand, an isolated or consecutive treatment of the two approaches makes their integration difficult and carries the risk that future teachers will reproduce teaching models that are far removed from social constructivism, where content is not articulated in learning situations connected to the ideas and culture of the students (García-Barros, 2016; Porlán et al., 2010).

Similarly, there is ongoing debate regarding the relative emphasis placed on disciplinary and didactic knowledge in teacher training curricula. In their analysis of the current regulation of Primary Education Teacher Degrees, Pro-Bueno et al. (2022) indicate that "it seems that legislators have placed greater emphasis on the learning of scientific knowledge", and add that "while acknowledging its importance, the curricular proposal is clearly improvable as it prioritises what a teacher should know about science over how to teach it" (p.192). From our point of view, the same conclusion can be drawn when examining the mathematical competences established for future teachers by this law. Accordingly, inservice primary school teachers demand didactic-pedagogical training to a greater extent than strictly disciplinary training (Manso and Garrido-Martos, 2021), despite the

aforementioned problems with their scientific-mathematical background.

Regarding the other two STEM disciplines, Technology and Engineering, the debate is also open. Defining technology and differentiating it from engineering is a complex issue (Cavanagh and Trotter, 2008) that varies from country to country (Yata et al., 2020). Some authors even suggest that the two disciplines are so closely linked that they should be taught together (Barak, 2013). However, there is a general consensus that, within the STEM framework, Engineering should focus on solving real-world problems by exploring solutions, developing prototypes, and testing their feasibility, while prioritising the processes and products created (Simarro and Couso, 2021). In fact, several methodologies that follow this approach have been proposed, such as *Design-based Learning* and *Design Thinking*.

On the other hand, Technology within STEM is considered to be the set of tools –often digital– that are used during scientific, mathematical, or engineering practices, or the product developed through these processes. In this respect, some authors state that Technology should not be considered a discipline and should not be included in the STEM acronym alongside Science, Mathematics, and Engineering (McComas and Burgin, 2020). In addition, Technology is not only specific to these fields, but also to social sciences, arts, or technical crafts (Costantino, 2018). With this complexity in mind, this study also analyses the presence of Engineering and Technology in the curricula of the Primary Education Teacher Degrees and assesses whether they are really aligned with their conception in STEM education.

Research questions

As a useful element of debate about curricular changes in the context of the LOMLOE, this work analyses the initial training of primary school teachers in the different STEM disciplines. Specifically, the following research questions regarding the compulsory training in the Spanish Primary Education Teacher Degrees are addressed:

- P.1. How many ECTS credits are assigned to scientific-mathematical areas and disciplines?
- P.2. Which is the predominant content: scientific-mathematical or didactic?

Received: 02-08-2022 Accepted: 02-04-2023

- P.3. Are Technology and Engineering approached from a STEM perspective?
- P.4. Are there differences in credits and content between public and private universities?
- Understanding the approaches and timing assigned to these areas and disciplines will help promote reflection to articulate curriculum plans that contribute to STEM teacher education.

Method

A cross-sectional statistical study has been conducted based on data collected from the curricula of Spanish universities.

Sample

The study sample comprises the Primary Education Teacher Degree programs at the 37 public universities and 22 private universities in Spain that offered this training during the 2020-2021 academic year.

The curricula were reviewed using the information available on the universities' websites. For this study, all compulsory subjects with STEM content were selected. This selection consisted of 342 subjects, 225 from public universities and 117 from private universities.

The teaching guides for these subjects were then analysed, which include information on scopes, content, competences, methodology, timing, and assessment.

Procedure

To answer the *first two research questions*, the contents and timing of the subjects in the sample were identified based on the teaching guides. According to the most frequent division, the contents were classified into: Mathematics, Physics, Chemistry, Biology, Geology, Didactics of Mathematics, and Didactics of Science.

For the calculation of credits assigned to the contents, the timing indicated in the teaching guides was considered. If this was not explicitly

Received: 02-08-2022 Accepted: 02-04-2023

identified, the credits were distributed proportionally to the number of content descriptors.

There are subjects solely dedicated to mathematical or scientific content (sometimes incorporating more than one scientific discipline), subjects dealing entirely with their didactics (curriculum, epistemology, methodologies and strategies, resources, assessment, etc.), and mixed subjects, where the credit count follows the aforementioned guidelines. Some subjects about didactics make use of certain scientific-mathematical concepts and models, but the ultimate aim is to illustrate methodological or didactic approaches. In such cases, credits were linked to didactics.

There are some contents whose categorisation to a certain discipline is not always obvious. In such cases, we followed a criterion consistent with the Primary Education curriculum. For instance, the specific contents of Astronomy (Universe, Solar System, Moon, etc.) were not considered in the analysis, as they are part of the Social Sciences in Primary Education and, therefore, in many universities they are also included in subjects about Geography, a discipline that has been excluded in this study and in the STEM disciplines. Conversely, regarding contents about nutrition and health, which are included in the Natural Sciences curriculum in Primary Education, they were assigned to Biology as it is the most related discipline of those established for the analysis.

The allocation of credits was carried out independently by two members of the research team. Discrepancies were discussed and resolved by agreement with the intermediation of a third researcher. Finally, the results were presented using descriptive statistical parameters.

To analyse the difference in the credits assigned to disciplinary and didactic content in each university, the parameter ε was defined, both for mathematics (ε_M) and science (ε_S), as follows:

$$\varepsilon = \frac{CR_{DID} - CR_{DIS}}{CR_{DID} + CR_{DIS}} \qquad (1)$$

Here, ε represents the relative difference between the credits assigned to didactics (CR_{DID}) and to disciplinary content (CR_{DIS}). In particular, $\varepsilon = 0$ indicates an equal distribution between the two, while $\varepsilon > 0$ reveals more emphasis on didactics and $\varepsilon < 0$ on disciplinary content. Extreme values correspond to universities that allocate all their credits to didactic content ($\varepsilon = 1$) or disciplinary content ($\varepsilon = -1$). In addition, we conducted a cluster analysis of these parameters to explore potential groups of

universities based on the prevalence of disciplinary or didactic content. The analysis aims to identify groups that are as homogeneous as possible while being different from each other. Various distance algorithms and group formation methods were tested to select those that made sense for the research. In particular, the Euclidean distance method and clustering using the Cutree algorithm (R Core Team, 2018) were chosen.

To address the *third research question*, we reviewed the contents of teaching guides for compulsory subjects dedicated specifically to Technology, analysing their coherence with a STEM approach (Moore et al., 2014). In addition, we checked for explicit references to STEM and specific Engineering content in the overall curriculum (Simarro and Couso, 2021).

Finally, to answer the *fourth research question*, which has a cross-sectional nature, and to make a comparison between public and private universities, the non-parametric Mann-Whitney U test was carried out to determine whether there are statistically significant differences ($p < 0.05$) in the distribution of credits. This choice was based on the previous application of the Shapiro-Wilk test, which allowed us to rule out the normality of the distributions.

For all research questions, calculation of statistical parameters, cluster analysis and hypotheses tests were carried out using the open-source software R (R Core Team, 2018).

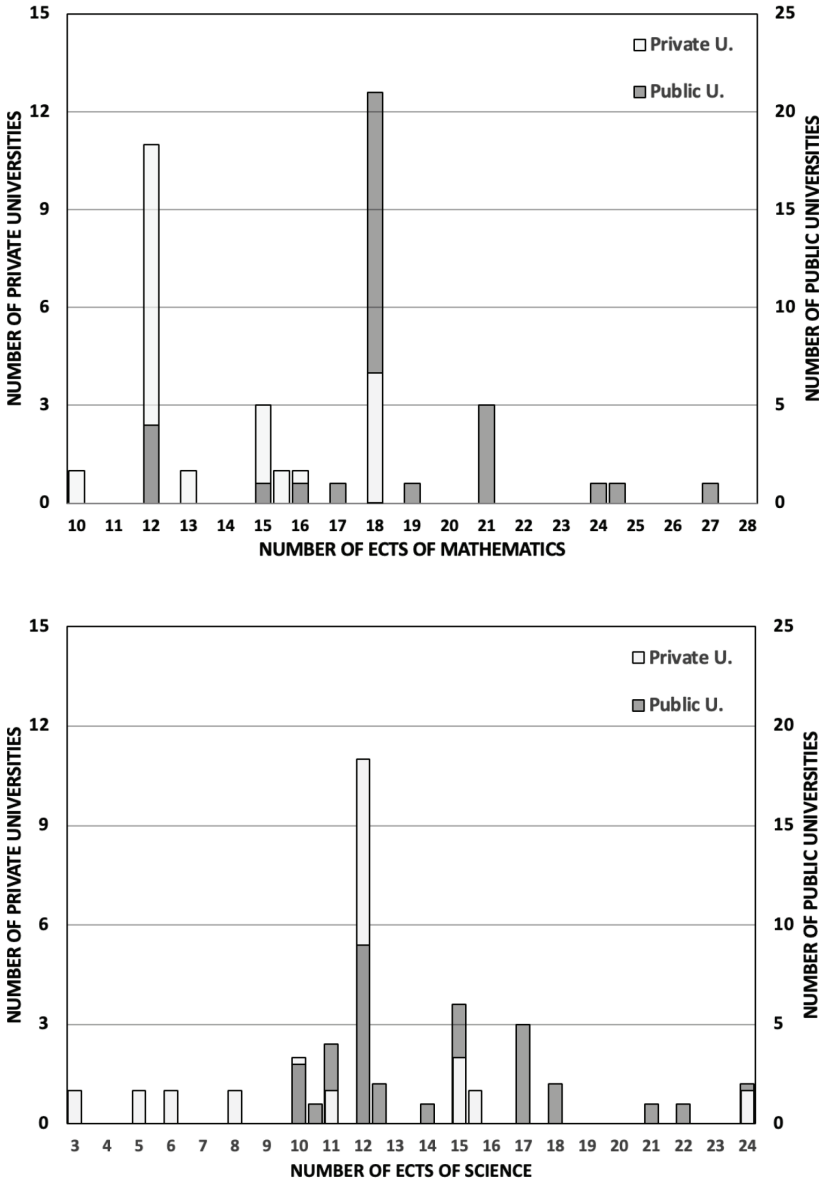
Results

The results are presented in three sections following the first three research questions. The last question is addressed across all sections.

Distribution of credits in scientific-mathematical disciplines

The total ECTS credits allocated to Mathematics and Science in the curricula (including their disciplinary and didactic contents) are shown in Graph I, and the descriptive statistics in Table I. Firstly, the dispersion of the results is striking, with some universities allocating little more than 10 credits to Mathematics and 5 to Science, while others allocate up to 25 credits to each. This dispersion is greater in Science than in Mathematics ($\sigma = 4.2$ vs. $\sigma = 3.6$). Secondly, there is a significant difference between

GRAPH I. Distribution of ECTS credits allocated to Mathematics (top) and Science (bottom), including their disciplinary and didactic contents, in public and private universities



Source: Compiled by the authors based on the study plans.

Received: 02-08-2022 Accepted: 02-04-2023

TABLE I. ECTS credits allocated to Mathematics and Science, including their disciplinary and didactic contents

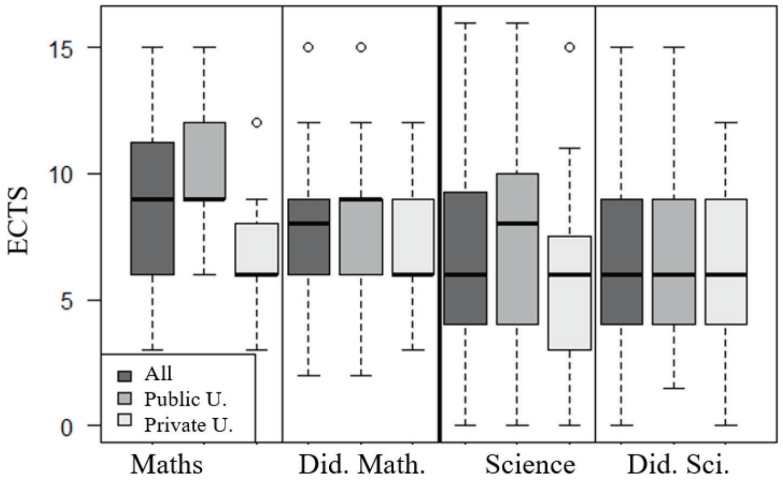
	PUBLIC UNIVERSITIES		PRIVATE UNIVERSITIES		ALL THE UNIVERSITIES	
	Mean	**Standard deviation**	**Mean**	**Standard deviation**	**Mean**	**Standard deviation**
Mathematics	18.2	3.2	13.8	2.5	16.6	3.6
Science	14.4	3.8	11.6	4.1	13.4	4.2
Mathematics + Science	32.6	4.6	25.4	5.9	29.9	6.2

universities, with public universities allocating 31.9% more credits to Mathematics (U = 694.0, p = 2.40e-06) and 24.1% more to Science (U = 559.5, p = 0.0149).

Graph II illustrates the distribution of ECTS credits assigned to disciplinary and didactic content for Mathematics and Science in both types of universities. The results indicate a statistically significant difference in the disciplinary credits for Mathematics (U = 681.5, p = 1.04e-05) and, although relevant, not significant for Science (U = 531.5, p = 0.0510), with public universities assigning more credits to these subjects. Therefore, the fundamental difference between universities is essentially due to the difference in time allocated to disciplinary content. Regarding the ECTS assigned to didactics, the difference between types of universities is not significant.

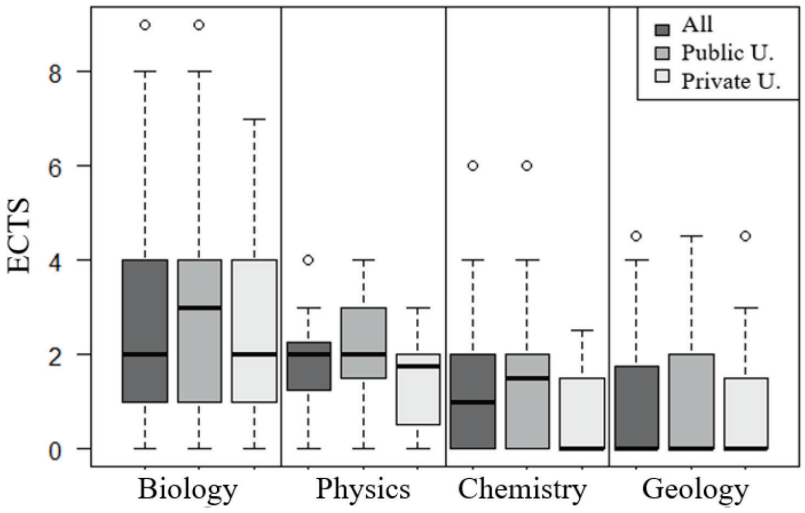
Finally, Graph III shows the results divided by disciplinary content: Biology, Physics, Chemistry, and Geology. The difference is evident, with the highest weight for Biology, followed by Physics, and quite low for Chemistry and Geology. This gradation aligns with the weight of these disciplines in the Primary Education curriculum. The disparity between types of universities is again evident, with public universities having a greater teaching load. On the other hand, it is remarkable that at least 50% of both public and private universities do not devote any time at all to Geology, nor to Chemistry in the case of private universities. There are no significant differences between the types of universities in any discipline.

GRAPH II. ECTS credits allocated to Mathematics, Science (including Biology, Physics, Chemistry, and Geology) and their respective Didactics, in public and private universities



Source: Compiled by the authors based on the teaching guides.

GRAPH III. ECTS credits allocated to Biology, Physics, Chemistry, and Geology in public and private universities



Source: Compiled by the authors based on the teaching guides.

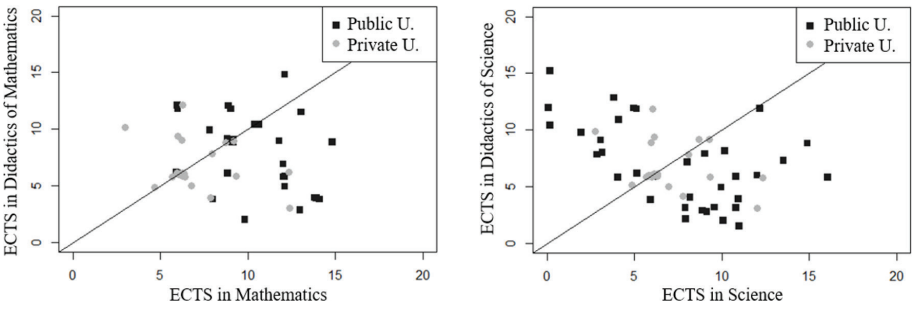
Received: 02-08-2022 Accepted: 02-04-2023

Scientific-mathematical training vs. didactic training

Firstly, the number of credits assigned by each university to scientific-mathematical disciplinary content and to didactic content is analysed. Graph IV displays Mathematics and Science separately; the straight line represents an equal distribution between both contents, with universities with more didactic credits at the top and those with more disciplinary credits at the bottom. Once again, a wide dispersion in teaching loads can be observed, with a greater dispersion in Science. When comparing disciplinary and didactic credits in public universities, a statistically significant difference is obtained in Mathematics ($U = 930.0$, $p = 0.0062$), with a predominance of disciplinary content, but not in Science ($U = 740.0$, $p = 0.5506$). In private universities, there are no significant differences between the two domains.

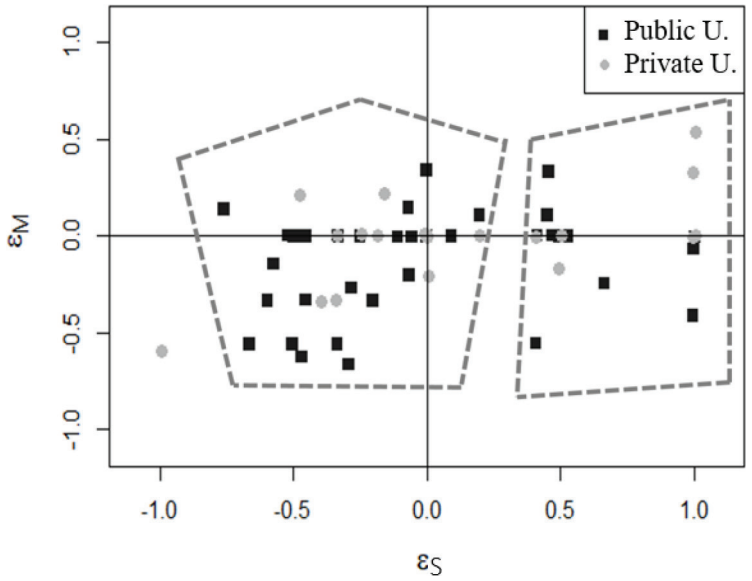
To illustrate the dominant formative role of each institution, Graph V shows jointly the results of the ε parameter for Mathematics and Science, while Table II provides a distribution of the parameter values. In Mathematics, there are no extreme cases with 0 ECTS for either content. The most common situation in both types of universities exhibits an equal distribution between disciplinary and didactic content ($\varepsilon_M = 0$), although there are many public universities where the more disciplinary approach dominates ($\varepsilon_M < 0$). On the other hand, there are more contrasting trends in Science, as most universities opt for one approach over the other,

GRAPH IV. ECTS credits allocated in each university to didactic versus disciplinary content, in Mathematics (left) and Science (right)



Source: Compiled by the authors based on the teaching guides.

GRAPH V. Parameter ε for Mathematics and Science of the universities analysed. The groups identified in the cluster analysis are indicated



Source: compiled by the authors.

TABLE II. Distribution of the values of the parameter ε for Mathematics and Science in the universities analysed

	MATHEMATICS (ε_M)			SCIENCE (ε_S)		
	<0	**0**	**>0**	**<0**	**0**	**>0**
Public Universities	15 (40.5%)	16 (43.2%)	6 (16.2%)	21 (56.8%)	3 (8.1%)	13 (35.1%)
Private Universities	5 (22.7%)	13 (59.1%)	4 (18.2%)	8 (36.4%)	6 (27.3%)	8 (36.4%)

Source: Compiled by the authors

with the disciplinary content approach dominating ($\varepsilon_s < 0$). The case of the Universitat Abat Oliba CEU is noteworthy, as training is exclusively oriented towards disciplinary content ($\varepsilon_s = -1$). Conversely, in institutions such as Alfonso X el Sabio, Autònoma de Barcelona, La Laguna, Mondragón, Oberta de Catalunya, Vic, and Zaragoza, training is entirely

Received: 02-08-2022 Accepted: 02-04-2023

centred on didactics ($\varepsilon_s = 1$). It is important to note that this last approach does not necessarily imply a lack of attention to disciplinary content in the subjects, which could be approached from the perspective of teaching and learning (school models, evolution of students' previous ideas, design of inquiry-based activities, etc.).

Graph V also shows the cluster analysis, which divides the sample into two major groups: $\varepsilon_s < 0.3$, comprising 26 public and 14 private universities (70.3% and 63.6%, respectively), and $\varepsilon_s > 0.3$, comprising 11 public and 7 private universities (29.7% and 31.8%, respectively). As a third group, Universitat Abat Oliba CEU appears separately due to its distinct approach. Therefore, the differentiation of university profiles is based on the model assumed for Science and its Didactics (horizontal axis), rather than on the model for Mathematics and its Didactics (vertical axis). While most universities prioritise explicit training in the disciplinary content of Science, a smaller group proposes to structure teacher training around its didactics.

To complement this view, Table III distributes the 298 Mathematics and Science subjects according to whether they address disciplinary or didactic content, or both. As such, it can be observed that public universities predominantly adopt a mixed approach for both disciplines, while private universities exhibit greater diversity.

Another interesting aspect to analyse is whether disciplinary training comes before, begins in the same year as didactics, or comes after it. The results are shown in Table IV, excluding the universities already mentioned which only teach disciplinary or didactic content in the case of Science. The data reveal that, in both Mathematics and Science, very

TABLE III. Subjects with a disciplinary-only, didactic-only, or mixed approach to content

	MATHEMATICS			SCIENCE		
	Disciplinary contents	**Mixed contents**	**Didactic contents**	**Disciplinary contents**	**Mixed contents**	**Didactic contents**
Public Universities	18 (16.8%)	74 (69.2%)	15 (14.0%)	21 (22.8%)	43 (46.7%)	28 (30.4%)
Private Universities	14 (26.4%)	23 (43.4%)	16 (30.2%)	15 (32.6%)	13 (28.3%)	18 (39.1%)

Source: Compiled by the authors based on the teaching guides.

TABLE IV. Distribution of the universities according to whether disciplinary training is prior to (G1), begins in the same year (G2) or is subsequent (G3) to didactic training

	MATHEMATICS			SCIENCE		
	G1	**G2**	**G3**	**G1**	**G2**	**G3**
Public Universities	14 (37.8%)	22 (59.5%)	1 (2.7%)	10 (27.0%)	22 (59.5%)	2 (5.4%)
Private Universities	9 (40.9%)	13 (59.1%)	0 (0.0%)	10 (45.5%)	8 (36.4%)	0 (0.0%)

Source: Compiled by the authors based on the study plans.

few universities teach didactics before disciplinary content. The majority of universities start teaching both in the same academic year.

Finally, Graph VI shows the distribution of Science and Mathematics subjects throughout the Primary Education Teacher Degree. It is shown that most of them are concentrated in the 2nd and 3rd years, with fewer subjects offered in the first year (which typically focuses on the basic training module) and in the fourth year (which is largely devoted to electives). No relevant differences were found between types of universities in this part of the analysis.

Technology and Engineering in the Primary Education Teacher Degree

Out of the total of 59 universities analysed, 44 of them have a specific compulsory subject whose title or contents mainly focus on the field of Technology. In almost all cases, these subjects carry 6 ECTS credits (88.6%). Their year distribution is as follows: 28 (63.6%) are taken in the first year, 13 (29.5%) in the second year, and 3 (6.8%) in the third or fourth year.

To analyse the role of Technology in these degrees, the contents declared in the teaching guides of these 44 subjects were reviewed. In this case, no relevant differences were found between public and private universities, so the results were grouped together.

As shown in Table V, the contents related to search for information and use of Internet resources, as well as those oriented towards the design of

GRAPH VI. Distribution of the universities based on when they begin (vertical) and end (horizontal) teaching the contents of the disciplines indicated

PUBLIC UNIVERSITIES				
Mathematics				
Year	**1º**	**2º**	**3º**	**4º**
1º	2	6	7	1
2º		3	10	8
3º			0	0
4º				0
Didactics of Mathematics				
Year	**1º**	**2º**	**3º**	**4º**
1º	0	1	6	2
2º		3	12	6
3º			3	3
4º				1
Science				
Year	**1º**	**2º**	**3º**	**4º**
1º	2	1	1	0
2º		9	7	1
3º			7	6
4º				0
Didactics of Science				
Year	**1º**	**2º**	**3º**	**4º**
1º	0	0	1	0
2º		2	9	5
3º			12	6
4º				2

PRIVATE UNIVERSITIES				
Mathematics				
Year	**1º**	**2º**	**3º**	**4º**
1º	3	1	1	1
2º		9	4	0
3º			1	2
4º				0
Didactics of Mathematics				
Year	**1º**	**2º**	**3º**	**4º**
1º	0	1	0	1
2º		6	5	0
3º			7	1
4º				1
Science				
Year	**1º**	**2º**	**3º**	**4º**
1º	3	1	0	0
2º		3	3	0
3º			6	2
4º				0
Didactics of Science				
Year	**1º**	**2º**	**3º**	**4º**
1º	0	0	0	0
2º		2	4	0
3º			9	3
4º				3

Source: Compiled by the authors based on the study plans.

activities, are particularly remarkable. In addition, various tools (videos, virtual reality, video games, simulations, virtual laboratories, etc.) are addressed, sometimes in relation to some active methodologies (inquiry, project-based learning, gamification, etc.) or integrative approaches (such as Science, Technology and Society, STS).

TABLE V. Frequency of the contents declared in the teaching guides of the subjects dedicated to Technology

Contents	Frequency		Contents	Frequency
Internet	44 (100.0%)		STS	9 (20.5%)
Web Resources	43 (97.8%)		SCORM	9 (20.5%)
Methodologies	24 (55.5%)		Programming	7 (15.9%)
Design of Activities	20 (45.5%)		Virtual Reality	6 (13.6%)
Curriculum	15 (34.1%)		Robotics	5 (11.3%)
Videos	12 (27.3%)		Video Games	4 (9.1%)
Epistemology	11 (25.0%)		Virtual Laboratories	3 (6.8%)
VLE	11 (25.0%)		Scratch	2 (4.5%)
Assessment	10 (22.7%)		Simulations	2 (4.5%)

Source: Compiled by the authors based on the teaching guides.

Virtual Learning Environments (VLE) and e-Learning concepts (such as SCORM, an acronym for *Shareable Content Object Reference Model*, a standard for standardising virtual learning content) also appear in a relevant way in Table V. Thereafter it is concluded that the concept of technology that future teachers may acquire is closely linked to that of Information and Communication Technologies (ICT), although some new elements related to digitalisation are beginning to appear.

Finally, Table V includes some contents related to Technology in STEM education, such as robotics or programming, albeit to a lesser extent. These contents appear similarly in public universities, such as A Coruña, Girona, Illes Balears, and Rey Juan Carlos, and other private universities, such as Ramón Llull, Internacional de Catalunya, or Vic. It is worth mentioning the unique case of the European University of Madrid, which has a 4 ECTS subject called "Introduction to Programming".

Regarding Engineering, no references to specific concepts of this discipline, such as engineering design, prototyping, or computational thinking, were found in the set of compulsory subjects analysed, except for the aforementioned subject at the European University. Accordingly, specific methodologies such as *Design-based Learning* or *Design Thinking* are

absent in the teaching guides. It is therefore clearly concluded that this discipline is not included in the compulsory training of future teachers.

Finally, it is worth noting that, surprisingly, references to STEM or STEAM have only been found in the teaching guides of two private universities: Internacional de La Rioja and Mondragón. In fact, in the latter there is a subject entitled "STEAM in Primary Education", although it seems to include content related to the teaching of language and mathematics and the use of associated technological tools.

Discussion

The aim of this study is to analyse STEM education in Primary Education Teacher Degrees in Spain (in terms of timing and approaches) as a point of debate about its adequacy for achieving STEM education goals. Firstly, it has been found that there is no STEM training understood in a global way (Moore et al., 2014). There are no compulsory subjects related to this approach, and there is no mention of it in the teaching guides for Science, Mathematics and Technology subjects either, with a few exceptions. This result contradicts the objectives of the LOMLOE and is surprising, given the ubiquity of the acronym in the specialised literature (Toma and García-Carmona, 2021), and reflects the slowness or resistance to transferring the results of educational research to the classroom, also at university (Carr et al., 2012).

Analysing each discipline, the findings are similar in Mathematics and Science in terms of the most relevant aspects. Firstly, there is a wide dispersion in the ECTS credits allocated to these disciplines, which agrees with other studies on public universities (Castro-Rodríguez and Montoro, 2021; García-Barros, 2016; Naya-Riveiro et al., 2021) and private universities (Nolla et al., 2021). Furthermore, it has been found that the difference between types of universities is essentially due to the lesser amount of time dedicated to disciplinary content in private universities, and not so much to the time spent in didactic content. In general, these values are below the recommendations of the *White Paper* on Primary Education Teacher Degrees (ANECA, 2005, p.201), which suggests a minimum of 20 ECTS for Mathematics (Naya-Riveiro et al., 2021) and a higher teaching load than those obtained for Science. In fact, this poor training is associated with the lack of confidence and insecurity that primary school

teachers experience when teaching these disciplines (García-Ruiz and Sánchez, 2006; Jarvis and Pell, 2004).

Another issue for debate in the degree reforms will be the distribution of credits among scientific disciplines, especially when this load can be distributed among different departments. Here it has been observed that the contents of the teaching guides reflect a distribution consistent with the presence of these disciplines in the Primary Education curriculum, with greater weight given to Biology, and then to Physics. On the other hand, there is a scarce presence of Geology and, especially in private universities, of Chemistry. The scientific content covered in each discipline can be consulted in García-Barros (2016).

Another important decision is whether to focus the training more on disciplinary content or on didactic content. It has been found that in most universities there is an equal distribution between both domains in Mathematics. Still, where this is not the case, it is almost always due to an excess of disciplinary content, which is much more pronounced in public universities. The result is similar to that of other studies (Naya-Riveiro et al., 2021), although this work provides the differentiating nuance of identifying the specific content and its ECTS timing. In Science, public universities seem to be more polarised between the two domains, with a greater emphasis on disciplinary content, while private universities maintain a better balance between the two. In this respect, Cortés et al. (2012) found that, at the beginning of their training, primary preservice teachers place greater importance on disciplinary content; however, as they progress through the degree program (and complete field experiences), the value given to didactic content increases.

Beyond content, there is a rich controversy between those who advocate for disciplinary and didactic content to be addressed consecutively or simultaneously. On the one hand, it is argued that deficient disciplinary training can hinder the construction of PCK (Cañal, 2008; Martínez-Borreguero, 2022) and that the relevance of students' prior ideas and conceptual errors is not sufficiently addressed (Porlán et al., 2010). However, consecutive approaches can make it difficult to integrate bot domains (García-Barros, 2016) and design contextualised learning situations that activate and mobilise essential knowledge, enabling the adoption of future teaching roles and promoting metacognition (Dignath and Büttner, 2008). Our study has revealed that Mathematics has a higher prevalence of subjects in which both disciplinary and didactic contents

Received: 02-08-2022 Accepted: 02-04-2023

are taught in an interrelated manner. This is also the case in Science, although less markedly so, which has been severely criticised (García-Barros, 2016). In both cases, didactic content is hardly ever taught before disciplinary content. In fact, there are only a few universities where either subject is taught in the final year, which could help to link this knowledge with the final field experiences, where students can really put into practice their previous learning in the degree and integrate both disciplinary and didactic domains (Naya-Riveiro et al., 2021).

With regards to Engineering, our study has shown that it is left out of the curricula of the Primary Education Teacher Degrees, as has been noted by other authors (Castro-Rodríguez and Montoro, 2021). There are no concepts such as engineering design, process optimisation, or computational thinking. Nor are there any references to methodologies associated with this approach, such as *Design-based Learning* or *Design Thinking*. This absence is not surprising, since this type of content was included for the first time in the Spanish Primary Education curriculum with the recent LOMLOE. Even so, this does not guarantee its rapid and effective inclusion in the classroom. In fact, in other countries, the transfer of engineering content to the classroom has proven to be costly, despite curricular support (Carr et al., 2012).

Finally, Technology is addressed in most degree programs through specific compulsory subjects. The approach for Technology is generally focused on ICT (search for information and resources, digital tools, etc.), which seems coherent with the specific competences established in the *White Paper* on Primary Education Teacher Degrees, i.e., "Ability to use and incorporate information and communication technologies appropriately in teaching-learning activities" (ANECA, 2005, p.90). Even so, following Gewerc and Montero (2015), it can be criticised that the notion of technology is limited to an instrumental perspective. In fact, in the STEM approach, technology is not exclusively associated with ICT, but is more of a tool or result of solving real-world problems that are viewed from a global perspective in Mathematics, Science, and/or Engineering. In this regard, some teaching guides –typically related to Physics– refer to technological devices associated with simple machines or electromagnets, but once again, not from a true STEM perspective. From our perspective, it would be advisable to genuinely interrelate Engineering and Technology with Science and Mathematics, to address the design and construction of technological devices that solve real-world problems and satisfy

social needs: waste separation, medical applications, space exploration, etc. (Simarro and Couso, 2021).

Conclusions

This study examines jointly the initial training of primary school teachers in STEM disciplines in both public and private universities. From this research, we can conclude that nowadays undergraduate programs do not include STEM training with a global approach in their compulsory itineraries. Engineering is absent from the curricula, and Technology is usually approached from an ICT-centred perspective, far removed from its conception in STEM education.

In addition, the number of ECTS credits in Science and Mathematics is below the recommendations established in the *White Paper* on Primary Education Teacher Degrees (ANECA, 2005, p.201), particularly in private universities. However, their relative weight does seem to be in line with the teaching load of the disciplines in Primary Education (Mathematics, Biology, Physics, Chemistry, and Geology, in that order).

Furthermore, there are important differences between the relevance given by universities to disciplinary and didactic content, especially in the case of Science, which is a very hot debate. In this line, it can be seen that there is a predominance of subjects that simultaneously combine both contents, although this interrelation is weaker in Science than in Mathematics.

This study is highly relevant at this moment given the renewed importance of the so-called STEM competence since the approval of the LOMLOE. This context requires an update in the curricula of the Primary Education Teacher Degrees, which Spanish universities will soon have to address. For this purpose, the review, analysis, and comparison presented in this study are undoubtedly valuable and highlight the need to specifically include Technology and Engineering aspects with a STEM perspective, without compromising the time devoted to Mathematics and Science, which is already less than recommended.

The limitations of the study presented may arise from the fact that information on the contents and their timing in ECTS credits was derived from the teaching guides, and not from the real activities implemented by the trainers. This can lead to biases in the results, due to tensions

Received: 02-08-2022 Accepted: 02-04-2023

between planned and executed training (Porlán et al., 2010). Moreover, like other research approaches, analysing teaching guides involves adopting criteria that may differ from the actual timing of STEM training content. In addition, the analysis was restricted to compulsory training, ignoring elective subjects which, although they are not linked to the general profile of graduates, provide a suitable space to focus on the didactic aspects of STEM disciplines (Naya-Riveiro et al., 2021) and their interaction during problem-solving.

For this reason, in future work, we propose to study elective subjects by analysing the teaching guides related to STEM. Another future line of research could be to investigate whether the conclusions presented here hold true for Double Degrees, whose variety is wide and where it is not always easy to combine the syllabuses of the degrees involved. Surveys to teacher trainers would also provide a complementary view that would undoubtedly contribute to the debate on the best way to reform the curricula of the Primary Education Teacher Degrees.

Acknowledgements

To Irene González Muñoz and Cristina Rodríguez Mayordomo, beneficiaries of a collaboration grant from the Spanish Ministry of Education and Vocational Training during the 2020-2021 academic year, for their assistance in data collection for this study.

Bibliographical references

Aguilera, D., & Ortiz-Revilla, J. (2021). STEM vs. STEAM education and student creativity: A systematic literature review. *Education Sciences, 11*, 331. https://doi.org/10.3390/educsci11070331

Amat, A., Martínez-Chico, M., & Jiménez-Liso, M. R. (2022). Formación de maestras por implementación de secuencias en su propio contexto de aula: red sistémica para el análisis de las entrevistas pre-post. [Teacher training by implementing instructional sequence in Primary School teachers' classroom context: systemic network for analyzing pre-post interviews]. *Revista Interuniversitaria de Formación del Profesorado, 97*(36.1), 35–56. https://doi.org/10.47553/rifop.v97i36.1.91928

ANECA(2005).*LibroBlanco.TítulodeGradoenMagisterio.Volumen1*.Agencia Nacional de Evaluación de la Calidad y Acreditación. https://www.aneca.es/documents/20123/63950/libroblanco_jun05_magisterio1.pdf/bd7fdceb-075e-6256-b769-f89502fec8aa?t=1654601800472

Appleton, K. (2003) How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in Science Education, 33*(1), 1–25. https://doi.org/10.1023/A:1023666618800

Barak, M. (2013). Teaching engineering and technology: cognitive, knowledge and problem-solving taxonomies. *Journal of Engineering, Design and Technology, 11*(3), 316–333. https://doi.org/10.1108/JEDT-04-2012-0020

Belbase, S., Mainali, B. R., Kasemsukpipat, W., Tairab, H., Gochoo, M., & Jarrah, A. (2021). At the dawn of science, technology, engineering, arts, and mathematics (STEAM) education: prospects, priorities, processes, and problems. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/1.1080/0020739X.2021.1922943

Brunsell, E. (2012). The engineering design process. En E. Brunsell (Ed.), *Integrating engineering and science in your classroom* (pp. 3–5). NSTA Press.

Cañal, P. (2008). ¿Cómo orientar la formación inicial del profesorado de primaria en didáctica de las ciencias experimentales? En M.R. Jiménez-Liso (Ed.), C*iencias para el mundo contemporáneo y formación del profesorado en Didáctica de las Ciencias Experimentales* (pp. 256–263). Universidad de Almería.

Cardetti, F., & Truxaw, M. P. (2014). Toward improving the mathematics preparation of elementary preservice teachers. *School Science and Mathematics, 114*(1), 1–9. https://doi.org/10.1111/ssm.12047

Carr, R. L., Bennett IV, L. D., & Strobel, J. (2012). Engineering in the K-12 STEM standards of the 50 U.S. states: An analysis of presence and extent. *Journal of Engineering Education, 101*(3), 539–564. https://doi.org/10.1002/j.2168-9830.2012.tb00061.x

Castro-Rodríguez, E., & Montoro, A. B. (2021). Educación STEM y formación del profesorado de Primaria en España. [STEM Education and Primary Teacher Training in Spain]. *Revista de Educación, 393*, 353–378. https://doi.org/10.4438/1988-592X-RE-2021-393-497

Received: 02-08-2022 Accepted: 02-04-2023

Cavanagh, S., & Trotter, A. (2008). Where's the "T" in STEM. *Education Week, 27*(30), 17–19.

Cortés, A. L., Gándara, M., Calvo, J. M., Martínez-Peña, B., Ibarra, J., Arlegui, J., & Gil-Quílez, M.J. (2012). Expectativas, necesidades y oportunidades de los maestros en formación ante la enseñanza de las ciencias en la Educación Primaria. [Expectations, needs and opportunities of pre-service teachers in view of science teaching in Primary Education]. *Enseñanza de las Ciencias, 30*(3), 155–176. https://doi.org/10.5565/rev/ec/v30n3.597

Costantino, T. (2018). STEAM by another name: Transdisciplinary practice in art and design education. *Arts Education Policy Review, 119*(2), 100–106. https://doi.org/10.1080/10632913.2017.1292973

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*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

García-Barros, S. (2016). Conocimiento científico Conocimiento didáctico. Una tensión permanente en la formación docente. [Science knowledge and Education knowledge. A continuous strain in teachers' education]. *Campo Abierto, 35*(1), 31–44. https://mascvuex.unex.es/revistas/index.php/campoabierto/article/view/2825

García-Ruiz, M., & Sánchez, B. (2006). Las actitudes relacionadas con las Ciencias Naturales y sus repercusiones en la práctica docente de profesores de primaria. *Perfiles Educativos, 114*, 61–89. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-26982006000400004&lng=es&nrm=iso

Gewerc, A., & Montero, L. (2015). Conocimiento profesional y competencia digital en la formación del profesorado. El caso del Grado de Maestro en Educación Primaria. [Professional knowledge and digital competency in teacher education. The case of Elementary Teacher Education Degree]. *Revista Latinoamericana de Tecnología Educativa, 14*(1), 31–43. https://doi.org/10.17398/1695-288X.14.1.31

INEE (2012). *TEDS-M. Estudio internacional sobre la formación en matemáticas de los maestros. Informe español*. Instituto Nacional de Evaluación Educativa. https://sede.educacion.gob.es/publiventa/d/15408/19/00

Imbernón, F., & Colén, M. T. (2014). Los vaivenes de la formación inicial del profesorado. Una reforma siempre inacabada. *Tendencias Pedagógicas, 24*, 265–284. https://revistas.uam.es/tendenciaspedagogicas/article/view/2106

Jarvis, T., & Pell, A. (2004). Primary teachers' changing attitudes and cognition during a two-year science in-service programme and their effect on pupils. *International Journal of Science Education, 26*(14), 1787–1811. https://doi.org/10.1080/0950069042000243763

Manso, J., & Garrido-Martos, R. (2021). Formación inicial y acceso a la profesión: qué demandan los docentes. [Initial training and access to the profession: teachers' demands]. *Revista de Educación, 393*, 293–319. https://doi.org/10.4438/1988-592X-RE-2021-393-494

Martínez-Borreguero, G., Naranjo, F. L., & Mateos, M. (2022). Cognitive and emotional development of STEM skills in primary school teacher training through practical work. *Education Sciences, 12*, 470. https://doi.org/10.3390/educsci12070470

McComas, W. F., & Burgin, S. R. (2020). A critique of "STEM" education. *Science & Education, 29*(4), 805–829. https://doi.org/10.1007/s11191-020-00138-2

MEFP (2022a). *24 propuestas de reforma para la mejora de la formación docente*. Ministerio de Educación y Formación Profesional. https//educagob.educacionyfp.gob.es/comunidad-educativa/profesorado/propuesta-reforma.html

MEFP (2022b). *Estadísticas de la Educación*. Ministerio de Educación y Formación Profesional. https://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas.html

Ministerio de la Presidencia (2021). *España 2050: Fundamentos y propuestas para una Estrategia Nacional de Largo Plazo.* https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/200521-Estrategia_Espana_2050.pdf

Moore, T. J., Glancy, A. W., Tank, K. M., Kersten, J. A., Smith, K. A, & Stohlmann, M. S. (2014). A framework for quality K-12 engineering education: research and development. *Journal of Pre-College Engineering Education Research, 4*(1), 2. https://doi.org/10.7771/2157-9288.1069

Naya-Riveiro, M. C., Gómez-Sánchez, T. F., Rumbo, M. B., & Segade, M. E. (2021). Estudio interregional comparado de la educación matemática en la formación inicial del profesorado de Educación Primaria. [Comparative interregional study of mathematical education for preservice primary teachers]. *RELIME. Revista Latinoamericana de Investigación en Matemática Educativa, 24*(2), 207–233. https://doi.org/10.12802/relime.21.2424

Nolla, Á., Muñoz, R., Cerisola, A., & Fernández, B. (2021). La formación inicial de los maestros en matemáticas y su didáctica. [Initial teacher training in elementary mathematics and its didactics]. *Revista Interuniversitaria de Formación del Profesorado, 96*(35.1), 185–208. https://doi.org/10.47553/rifop.v96i35.1.85882

Nortes Checa, A., & Nortes Martínez-Artero, R. (2013). Formación inicial de maestros: Un estudio en el dominio de las matemáticas. [Primary school teachers' initial training: A study in math domain]. *Profesorado, Revista de Currículum y Formación del Profesorado, 17*(3), 185–200. https://revistaseug.ugr.es/index.php/profesorado/article/view/19684

Porlán, R., Martín del Pozo, R., Rivero, A., Harres, J., Azcárate, P., & Pizzato, M. (2010). El cambio del profesorado de ciencias I: marco teórico y formativo. [Science teacher change I: Theoretical and teacher education framework]. *Enseñanza de las Ciencias, 28*(1), 31–46. https://doi.org/10.5565/rev/ensciencias.3619

Pro-Bueno, A., Pro-Chereguini, C., & Cantó, J. (2022). Cinco problemas en la formación de maestros y maestras para enseñar ciencias en Educación Primaria. [Five problems in training teachers to teach science in Primary Education]. *Revista Interuniversitaria de Formación del Profesorado, 97*(36.1), 185–202. https://doi.org/10.47553/rifop.v97i36.1.92510

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.Rproject.org/

Real Decreto 157/2022, de 1 de marzo, por el que se establecen la ordenación y las enseñanzas mínimas de la Educación Primaria. *Boletín Oficial del Estado, 52*, de 2 de marzo de 2022. https://www.boe.es/boe/dias/2022/03/02/pdfs/BOE-A-2022-3296.pdf

Sánchez-Urán, L. (2019). La formación inicial de los maestros de Educación Primaria en España: universidades y planes de estudios. En J. Manso (Ed.), *La formación inicial del profesorado en España* (pp. 12–30). Ministerio de Educación y Formación Profesional.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Simarro, C., & Couso, D. (2021). Engineering practices as a framework for STEM education: a proposal based on epistemic nuances. *International Journal of STEM Education, 8*, 53. https://doi.org/10.1186/s40594-021-00310-2

Socas, M. M. (2011). Aprendizaje y enseñanza de las Matemáticas en Educación Primaria. Buenas prácticas. *Educatio Siglo XXI, 29*(2), 199–224. https://revistas.um.es/educatio/article/view/133031

Toma, R. B., & García-Carmona, A. (2021). «De STEM nos gusta todo menos STEM». Análisis crítico de una tendencia educativa de moda. [«Of STEM we like everything but STEM». A critical analysis of a buzzing educational trend]. *Enseñanza de las Ciencias, 39*(1), 65–80. https://doi.org/10.5565/rev/ensciencias.3093

Toma, R. B., Greca, I. M., & Meneses-Villagrá, J. A. (2017). Dificultades de maestros en formación inicial para diseñar unidades didácticas siguiendo la metodología de indagación. [Elementary pre-service teachers' difficulties for designing science-teaching units by inquiry]. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 14*(2), 442–457. http://dx.doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2017.v14.i2.11

Yata, C., Ohtani, T., & Isobe, M. (2020). Conceptual framework of STEM based on Japanese subject principles. *International Journal of STEM Education, 7*(12). https://doi.org/10.1186/s40594-020-00205-8

Contact address: Germán Ros Magán. Universidad de Alcalá, Facultad de Educación, Dpto. de Física y Matemáticas. Avda. de Madrid 1, CP 19001, Guadalajara. E-mail: german.ros@uah.es

Received: 02-08-2022 Accepted: 02-04-2023

Effect of advanced high school major on mathematical performance at university: a comparative study in Business Administration degrees

Efecto de la especialidad en bachillerato en el rendimiento matemático en la universidad: un estudio comparativo en grados de Administración de Empresas

https://doi.org/10.4438/1988-592X-RE-2023-402-597

José Luis Arroyo-Barrigüete
https://orcid.org/0000-0002-3660-3933
Universidad Pontificia Comillas

Susana Carabias López
https://orcid.org/0003-1375-1308
Universidad Pontificia Comillas

Adolfo Hernández
https://orcid.org/003-1078-2328
Universidad Complutense de Madrid

Marina Segura
https://orcid.org/003-00002-6068-960X
Universidad Complutense de Madrid

Abstract

Previous studies suggest that the major taken at high school is a very relevant variable in predicting average marks during the first year of business administration (BA) degrees. However, the effect on particular subjects and

Received: 15-10-2022 Accepted: 30-05-2023

specifically on mathematics, is not fully solved, as there are very few studies on the topic, the evidence is contradictory and there are no comparative studies between different centers. This paper compares the results at two different universities: 873 students of the BA degree at the Universidad Complutense de Madrid and 822 at the Universidad Pontificia Comillas. The information was obtained from the institutional databases at the two centers, selecting from among the students enrolled between the academic years 2009/2010-2021/2022 those for whom complete information was available. Methodologically, the use of regression models has been combined with interpretable neural networks to ensure the robustness of the results. The analysis shows that, at both universities and independently of the methodological approach, the results are virtually identical: students from the science major outperform their peers from the social sciences major in business mathematics I and II. From the point of view of teaching practice, these results have two implications. Firstly, it seems necessary to reflect on the focus and content of mathematics in the social sciences major, seeking greater alignment with the requirements of the degrees linked to this major. Secondly, regarding teaching practice in the first year of university, it seems necessary to rethink teaching strategies in mathematics, bearing in mind the characteristics and learning styles of students from the social sciences major.

Keywords: Mathematics, higher education, academic performance, business degree, advanced high school.

Resumen

Estudios previos apuntan a que la especialidad cursada en bachillerato es una variable muy relevante en la predicción del rendimiento medio durante el primer curso en grados en Administración y Dirección de Empresas (ADE). Sin embargo, el efecto en asignaturas concretas y específicamente en las de matemáticas, no está plenamente resuelto, pues además de que existen muy pocos estudios al respecto, las evidencias son contradictorias y no existen estudios comparativos entre varios centros. En este trabajo se comparan los resultados en dos universidades diferentes: 873 alumnos del grado en ADE de la Universidad Complutense de Madrid y 822 de la Universidad Pontificia Comillas. La información se ha obtenido de las bases de datos institucionales de ambos centros, seleccionando de entre todos los alumnos matriculados entre los cursos 2009/2010-2021/2022 aquellos para los que se disponía de información completa. Metodológicamente se ha combinado el uso de modelos de regresión con redes neuronales interpretables, para asegurar la robustez de los resultados. El análisis muestra que, en ambas universidades e independientemente de la aproximación metodológica, los resultados son virtualmente idénticos: los alumnos procedentes del bachillerato de ciencias presentan un rendimiento

Received: 15-10-2022 Accepted: 30-05-2023

académico sensiblemente mejor en matemáticas empresariales I y II que sus compañeros de ciencias sociales. Desde el punto de vista de la práctica docente, estos resultados tienen dos implicaciones. En primer lugar, parece necesario llevar a cabo una reflexión sobre el enfoque y contenidos de matemáticas en el bachillerato de ciencias sociales, buscando un mayor alineamiento con los requerimientos de los grados ligados a esta especialidad. En segundo lugar, respecto a la práctica docente en el primer curso universitario, parece necesario repensar las estrategias didácticas en matemáticas, considerando las características y el estilo de aprendizaje de los alumnos procedentes del bachillerato de ciencias sociales.

Palabras clave: Matemáticas, educación superior, rendimiento académico, grado en administración de empresas, bachillerato.

Introduction

According to previous research (Arroyo-Barrigüete *et al*., 2020a), the major at advanced high school has a considerable impact on the academic performance of Business Administration (BA) students during the first year of their university studies. However, this result refers to average performance, not to specific subjects. In the particular case of mathematics subjects, there are very few studies and the evidence is also contradictory. Thus, for example, Gonzalez Veiga *et al*. (1999) concluded that, both for the former five-year BA degree and for the three-year BA diploma, students who had taken the science major in their advanced high school studies achieved a better performance in BA mathematics than their peers who majored in social sciences. This study was based on a sample of 242 undergraduates in the five-year degree and 169 undergraduate students in the three-year diploma, all at the University of Oviedo. Martínez de Ibarreta *et al*. (2010), working on a sample of 554 BA students at the Universidad Pontificia Comillas, concluded that having taken a high school major in science had a positive effect on performance in quantitative subjects, especially in mathematics. However, these results do not coincide with a more recent study by Dávila *et al*. (2015) for students of the BA degree in this subject, whose conclusions suggest that there are no differences between the two groups of students. In this case, the authors worked on a sample of 279 students from the University of Las Palmas de Gran Canaria.

Received: 15-10-2022 Accepted: 30-05-2023

Knowing that mathematics subjects in the first year of the BA degree tend to present the highest drop-out rate[1] is key to identifying those factors that influence this. Therefore, it is necessary to know the effect of a variable that, according to the aforementioned studies, could have a considerable impact.

This study has been carried out by university lecturers of quantitative methods in Business Administration studies, with the ultimate goal of improving teaching and learning processes, so that synergies between their teaching and research work can be generated (Jaworski, 1998). This research paper can be framed within the area of knowledge of Mathematics Education, whose philosophical and theoretical foundation is presented in Ernest (2016).

The need to accurately define the objective of the research and test out a concrete hypothesis can make it difficult to connect with the complex problem of improving teaching and learning processes. Godino *et al*. (2021) offer a solution to this difficulty that includes three contributions to research in Mathematics Education, based on the so-called onto-semiotic model (Godino and Batanero, 1994, Godino *et al*., 2007; Godino *et al*., 2019). The first of these contributions is the specification of the mathematical content of the problem under study. This specification is directly related to the name of the reference model, since it requires the development of two tasks: on the one hand, the ontological task of identifying the mathematical objects involved; on the other hand, the semiotic task of studying the relationships between these objects. The second contribution by Godino *et al*. (2021) is the decomposition of instructional processes into their different facets, which they call epistemic, cognitive, interactional, ecological, affective and mediational dimensions. Focusing on each of these facets allows such a complex problem as mathematics teaching processes to be efficiently analyzed; Hence, this is the approach chosen to draw the conclusions of this study. The third contribution, which consists of identifying criteria that contribute to good teaching practices through the so-called

[1] Annex 1 of this paper includes a cluster analysis, which shows that in the "Basic Training" subjects of the first year, mathematics exhibits behaviour that is clearly different from that of other subjects, while also presenting significantly higher levels of academic failure.

theory of didactic suitability, will also be considered. The concept of didactic suitability of a teaching process is defined as the degree to which it meets certain characteristics that allow it to be classified as optimal or adequate for its objectives. Suitability can be considered both in relation to the general teaching process and to a particular dimension.

Accordingly, the existence of discrepancies in previous studies and the importance of mathematics subjects in the academic failure of BA students justify this paper. In order to overcome certain limitations identified in previous studies, we have used significantly larger samples and, additionally, we have worked with two different universities - one public and the other private - whose students have a different sociodemographic profile.

Secondly, two types of causal models have been used - linear regression and neural networks. In this regard, the recent development carried out by Pizarroso *et al*. (2022) has paved the way for the use of neural networks for explanatory purposes, something that was not possible before. Some previous developments allowed an understanding of the functioning of the network to some extent, but this work allows a much simpler interpretation, which is relatively similar to that of a regression model. The enormous advantage is that a neural network model does not require an *a priori* functional specification, but the network itself is able to automatically detect any linear, non-linear or interaction effect. This avoids omitting relevant effects that could distort the result, and is therefore an extremely useful tool for validating the results of more conventional models: any discrepancy would indicate that some significant effect has been omitted[2].

This double validation, with large samples from two different universities and using different methodologies, will make it possible to verify to what extent the results are similar regardless of the university or methodological approach, which would confirm the robustness and external validity of the results.

[2] As this development is brand new, it is currently not possible to use the neural network model alone, since the authors of NeuralSens have not yet proposed hypothesis tests to determine when a variable is significant. This is why, in its current state, its main use in the social sciences is to validate the functional specification of conventional econometric models.

Method

Context

This study is part of a research project funded by the Spanish Ministry of Science and Innovation, whose objective is to develop algorithms to anticipate academic risk situations among university students. This work focuses on undergraduate studies in Business Administration, and on the effect of the high school major on academic performance in mathematics subjects taught in the first year of the degree. This choice is justified by the fact that the first year is key, which accounts for a large number of the dropouts, and mathematics subjects are in turn those with the highest level of academic failure.

Sample

The sample is composed of 1,695 students of the BA degree, of which 873[3] are students at the Universidad Complutense de Madrid (academic years from 2009/2010 to 2021/2022), and 822 at the Universidad Pontificia Comillas in the same period. The information was collected through the institutional databases of the two universities[4], obtaining data on all the students enrolled in the period considered. Subsequently, the data were cleaned, eliminating incomplete records, i.e. those for which no information was available on any of the variables required for the model. The study was carried out on all the students for whom complete information was available.

As seen in Table I, there are certain differences between the two groups of students, both in the percentage of women and in the percentage of students from the science major in advanced secondary education. Differences can also be seen in the university entrance marks and in the

[3] The initial sample consisted of a significantly larger number of students (see Annex 1), but many of them were eliminated from the study because there was no information available on the major studied at high school. The same happens in the sample from Universidad Pontificia Comillas, since there is no information on the high school major prior to the 2012/2013 academic year.

[4] Integrated Institutional Data System (Sistema Integrado de Datos Institucionales—SIDI) of the Universidad Complutense de Madrid (UCM), and General Secretary's Office of the Universidad Pontificia Comillas.

Received: 15-10-2022 Accepted: 30-05-2023

average marks in Business Mathematics I and II. Among other factors, this is due to a high percentage of students who do not take the first exam sitting at the first university (students who have been assigned a mark of zero). However, in the case of the second university, we hardly found any students that did not sit the exam. This induces a notable bias in the average mark, and is one of the main causes of the aforementioned difference in marks, in addition to, possibly, the different percentage of students from the science major. At any event, given that the objective is not to make a comparison of performance between the two universities, we will not delve further into this particular aspect, which furthermore presents considerable complexity.

TABLE I. Sample used in the analysis

	Sample size	% female	% science major	EvAU: average mark (sd)	Maths I: average mark (sd)	Maths II: average mark (sd)
U. Complutense de Madrid	873	43.99%	19.36%	7.02 (0.77)	2.75 (2.55)	3.83 (3.02)
U. Pontificia Comillas	822	50.24%	31.87%	7.71 (0.85)	5.33 (1.94)	5.73 (1.76)

Source: Compiled by author.

Procedure

The R programming environment (R Core Team, 2020) was used for data processing and model building, employing the packages *lmtest* (Zeileis and Hothorn, 2002), *lubridate* (Grolemund and Wickham, 2011), *lfe* (Gaure, 2013), *NbClust* (Charrad et al., 2014), *ggplot2* (Wickham, 2016), *car* (Fox and Weisberg, 2019), *RCurl* (Lang and CRAN team, 2019), *caret* (Kuhn, 2020), *gplots* (Warnes *et al*., 2020), *dplyr* (Wickham *et al*., 2022) and *NeuralSens* (Pizarroso *et al*., 2022).

In both samples, a linear regression model was adjusted, using the average mark in Business Mathematics I and II as a dependent variable, and the major taken at high school, the gender of the student, and his or her university entrance exam score (EvAU) as independent variables, the latter

Received: 15-10-2022 Accepted: 30-05-2023

being the equivalent of the SAT in the United States, since it is the score used in the university admission process. Due to problems of heteroscedasticity, robust estimations have been used, the absence of multicollinearity problems has been verified, and standardized variables have been used.

To confirm the functional specification validity, neural network models have also been adjusted using the same variables. Historically, neural network models were not suitable for the development of explanatory models due to their black-box nature. That is, they offered good predictions, in many cases better than those of more conventional econometric models, but it was not possible to interpret them, since the effect of each variable was unknown. However, several recent studies have solved this problem. Specifically, the development of Pizarroso *et al*. (2022), based on the *NeuralSens* algorithm, allows their interpretation in a very simple manner, since the slope of each variable is obtained, which would be the equivalent of the beta in a regression model. The nuance is that, unlike regression models, where the slope of each variable is unique (a single beta), in neural networks a slope is obtained for each piece of data. This implies that each variable will have a distribution of slopes. In the absence of non-linear effects, the distribution will be very narrow and its mean value will coincide exactly with the beta obtained in a regression model. The main advantage of using neural networks over regression models is that they do not require an *a priori* functional specification, and any non-linear effects in the data will be automatically identified by the network, without the need for the researcher to explicitly formulate them. Thus, if the results of the neural network coincide with those of the regression model, the functional specification of the second model would be confirmed to be correct. In the event of discrepancies, the regression model would be wrongly specified and would need to be modified, probably to include some non-linear effects not initially formulated.

The results section includes the three indicators proposed by Pizarroso *et al*. (2022) to interpret the results of a neural network: *mean sensitivity, sensitivity standard deviation* and *mean squared sensitivity*. As already indicated, in a neural network a sensitivity or slope is obtained for each piece of data, which is equivalent to a beta in a regression model, so that a given variable has as many betas as data in the sample. In other words, instead of obtaining a value for the slope of each variable, we obtain a density function, which in the absence of non-linear relationships will have a *mean sensitivity* equal to the slope of the regression model. The presence

Received: 15-10-2022 Accepted: 30-05-2023

of distributions with several modes or high dispersion (high values of *sensitivity standard deviation*) are clear indicators of the presence of non-linear effects or interactions. If detected, they would force the functional specification of the regression model to be modified[5]. Finally, the value of the *mean squared sensitivity* is the metric proposed by the authors to determine the relative importance of each of the predictors.

The main problem of neural networks is their propensity to overfitting, which is why taking certain precautions in the process of selecting their hyperparameters (number of neurons in the hidden layer and penalty or decay) is essential. In this study, we have opted for the net method, i.e. adjusting all the penalty combinations between 10^{-7} and 10^{-2}, and between 1 and 10 neurons in the hidden layer. In other words, 60 different neural networks have been adjusted, selecting the one with the lowest root mean square error (RMSE) as optimal. As a precaution, the 10-folds cross-validation procedure was applied in all cases to avoid over-fitting problems.

Finally, this analysis has been complemented with an equivalent one for the subject of Business Statistics, which, in addition to being the most similar subject in terms of content among all the basic-training subjects, presents the worst results (lower mean) of all the subjects except for Business Mathematics I and II (see Annex 1).

Results

Universidad Complutense de Madrid

The regression model obtained for the Universidad Complutense de Madrid students presents an R^2 of 0.18, and the three predictors have a positive and significant effect (alpha = 0.01). It can be seen that the major taken at high school has a considerable impact (see the left-hand side of Table II).

[5] The detection of these effects is very important since previous research has detected non-linear effects or interactions in mathematics performance. Thus, for example, the work of Arroyo-Barrigüete *et al.* (2020b) identified that among the students of BA-Law double degree, there was a group of students (around 30%), with a strong bias towards law, who used a rote learning approach to mathematics. This translated into an absence of correlation between performance in mathematics subjects, i.e. their performance in Mathematics II was not related to that obtained in Mathematics I, while, for the rest of the students the correlation was logically very strong. This type of possible interaction, which *a priori* are not easy to identify, are precisely those that the neural network model would allow us to detect.

As far as the neural network[6] is concerned (see the right-hand side of Table II), it does not seem to have detected non-linear effects or interactions, since the distributions are relatively narrow (*Sensitivity std. deviation* small, as shown in Table II) and unimodal (see bottom part of Figure I[7]). In fact, the coefficients estimated by the regression model are virtually identical to the mean sensitivity values obtained with the neural network ("Coef." and "Mean sensitivity" columns in Table II). This is an indicator of the adequacy of the functional specification used in the regression model.

TABLE II. Estimates obtained by the regression model and the neural network for students at the Universidad Complutense de Madrid (average mark in Business Mathematics I and II)

	Regression model (OLS)			**Neural network (NeuralSens)**		
	Coef.	**sd**	**p-value**	***Mean sensitivity***	***Sensitivity sd***	***Mean sq. sensitivity***
Constant	-0.25	0.05	8.6 E-8			
Gender (Female=1)	0.32	0.06	1.0 E-6	0.32	0.03	0.32
EvAU	0.28	0.03	< 2 E-16	0.29	0.02	0.29
High school major (Science=1)	0.54	0.08	2.0 E-10	0.56	0.05	0.56

Source: Compiled by author.

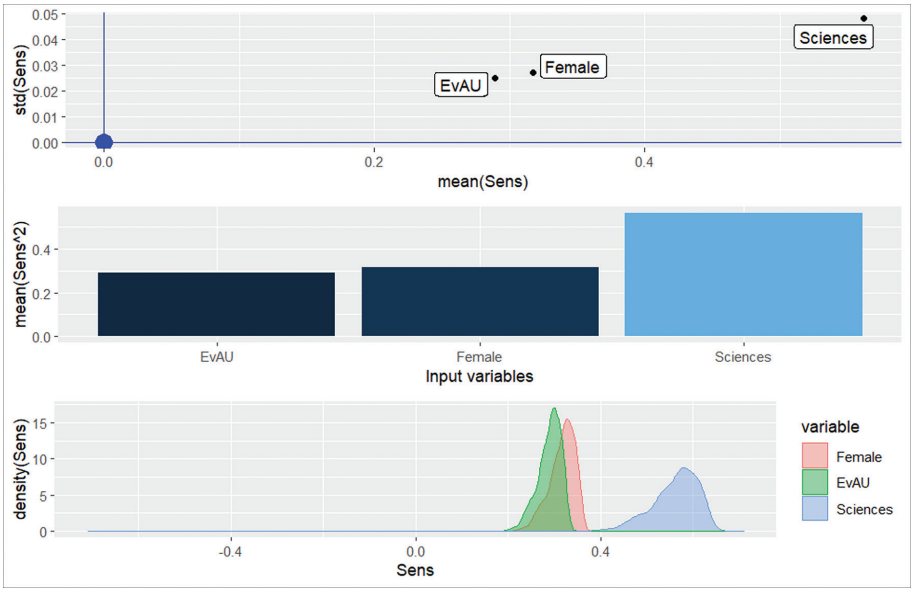
The consistency of both models confirms the importance of the variable under study: the high school major has a considerable effect on performance in Business Mathematics I and II, so that those students who take the science major have a significantly higher performance than their peers that take the social sciences major.

[6] The optimal neural network identified presents a 3-1-1 architecture, with a decay of 10^{-4}. The fact that we only find one neuron in the hidden layer is already a first indication that there are probably no complex non-linear effects.

[7] Figure I shows the results of the neural network sensitivity analysis performed with NeuralSens. It is divided into three parts: the lower window shows the distribution of the slopes (*sensitivity*) of each variable; in the middle window we find the mean squared sensitivity for each variable, which, as already indicated, allows us to determine the relative importance of each of the predictors; the upper part shows the *mean sensitivity* and the *sensitivity standard deviation* of each variable, which are precisely the values shown in Table II.

Received: 15-10-2022 Accepted: 30-05-2023

FIGURE I. Neural network results, including the metrics proposed by Pizarroso *et al.* (2022), for Universidad Complutense de Madrid (average mark in Business Mathematics I and II)



Source: Compiled by author.

In order to assess whether this effect extends to other related subjects, the analysis has been reproduced with the subject of Business Statistics, which is also taught during the first year and forms part of the "Basic Training" subjects which, after the two mathematics subjects, presents the worst results (lower mean) of all the subjects (see Annex 1). Given the focus of mathematics in the social sciences major, which has a greater content in statistics, it would be expected that students from this major would perform significantly better in this subject.

The results obtained are shown in Table III: on the left-hand side, the regression model, which obtains an R^2 of 0.08; on the right-hand side, the neural network (3-1-1 with a *decay* of 0.01). Once again, both models are consistent, since the coefficients estimated by the regression model are practically identical to the mean sensitivity values obtained with the neural network. The results indicate that there are no differences (at a 1% significance level) between students from the science and social sciences

major. However, at a significance level of 5%, there would be differences in favour of students from the science major. This result is even more striking than that obtained in the case of Business Mathematics I and II since, as already indicated, the social science major has a significantly higher statistics content than the science major.

TABLE III. Estimates obtained by the regression model and the neural network for students from the Universidad Complutense de Madrid (mark in Business Statistics)

	Regression model (OLS)			**Neural network (NeuralSens)**		
	Coef.	**sd**	**p-value**	***Mean sensitivity***	***Sensitivity sd***	***Mean sq. sensitivity***
Constant	-0.20	0.05	2.7 E-5			
Gender (Female=1)	0.38	0.07	2.6 E-8	0.38	0.16	0.41
EvAU	0.17	0.03	9.5 E-7	0.19	0.08	0.21
High school major (Science=1)	0.19	0.08	2.5 E-2	0.20	0.08	0.22

Source: Compiled by author.

Universidad Pontificia Comillas

The regression model obtained for the Universidad Pontificia Comillas students (left-hand side of Table IV) has an R^2 of 0.26, somewhat higher than in the previous sample, and the three predictors also have a positive and significant effect (alpha = 0.01). The optimal neural network identified (right-hand side of Table IV) presents a 3-1-1 architecture, with a decay of 0.01, somewhat lower than in the case of the Universidad Complutense de Madrid. It should be noted that the standard deviations of the sensitivities (*Sensitivity sd*) identified by *NeuralSens* are somewhat higher than in the case of the Universidad Complutense de Madrid, as can be seen in both Table IV and Figure II. This points to the possible existence of some non-linear effects in the EvAU and high school major variables[8], albeit not too pronounced.

[8] It might be appropriate to include a quadratic term for EvAU in the regression model. However, for the sake of comparability with the model of the Universidad Complutense de Madrid, it has been decided not to do so.

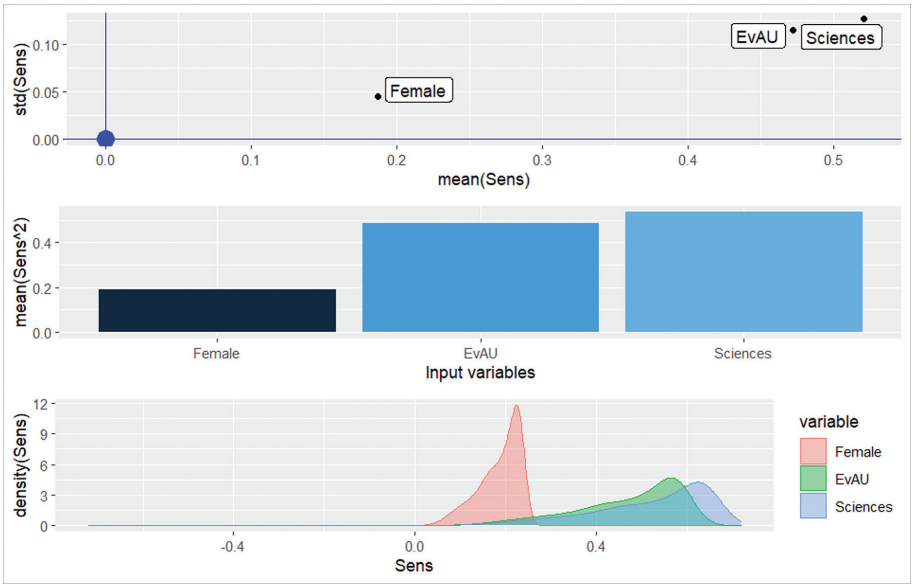
Received: 15-10-2022 Accepted: 30-05-2023

TABLE IV. Estimates obtained by the regression model and the neural network for students at Universidad Pontificia Comillas (average mark in Business Mathematics I and II)

	Regression model (OLS)			Neural network (NeuralSens)		
	Coef.	**sd**	**p-value**	***Mean sensitivity***	***Sensitivity sd***	***Mean sq. sensitivity***
Constant	-0.27	0.05	7.4 E-8			
Gender (Female=1)	0.19	0.06	1.8 E-3	0.19	0.05	0.19
EvAU	0.46	0.03	< 2 E-16	0.47	0.11	0.49
High school major (Science=1)	0.56	0.07	2.99 E-16	0.52	0.13	0.54

Source: Compiled by author.

FIGURE II. Neural network results, including the metrics proposed by Pizarroso *et al.* (2022), for Universidad Pontificia Comillas students (average mark in Business Mathematics I and II).



Source: Compiled by author.

Received: 15-10-2022 Accepted: 30-05-2023

In this case, the coefficients estimated by the regression model are again very similar to the mean values obtained with the neural network (*Mean sensitivity* column in Table IV). It is confirmed that, in this sample, the major taken in high school again has a considerable effect on performance in Business Mathematics I and II, with students from the science major performing better than those taking social sciences.

Reproducing this analysis for the subject of Business Statistics, we obtain the results shown in Table V (regression with an R^2 of 0.26 and a neural network of 3-1-1 with a decay of 0.01). We should indicate that although this exercise has been carried out for the purpose of completeness, in this case the comparison with the results obtained at the Universidad Complutense de Madrid is not appropriate for two reasons: at the Universidad Pontificia Comillas this subject is taken in the second year of the degree, which incorporates an uncontrolled effect due to the different maturity of the students, and this subject was incorporated into the curriculum in the 2015/2016 academic year, so the sample is somewhat reduced since there are no data for previous years, losing 428 records. At any event, the results of both causal models are, again, virtually identical to each other, and it is observed that the only relevant variable is the EvAU mark.

TABLE V. Estimates obtained by the regression model and the neural network for students at Universidad Pontificia Comillas (mark in Business Statistics)

	Regression model (OLS)			**Neural network (NeuralSens)**		
	Coef.	**sd**	**p-value**	***Mean sensitivity***	***Sensitivity sd***	***Mean sq. sensitivity***
Constant	-0.04	0.07	0.57			
Gender (Female=1)	0.00	0.09	0.96	0.00	0.00	0.00
EvAU	0.52	0.04	< 2 E-16	0.52	0.07	0.52
High school major (Science=1)	0.12	0.10	0.21	0.10	0.01	0.10

Source: Compiled by author.

Received: 15-10-2022 Accepted: 30-05-2023

Discussion and conclusions

This work is framed within the difficulty of the transition of students from high school to university, particularly with regard to mathematics subjects (Gueudet, 2008, De Guzman *et al.*, 1998). Specifically, we have studied the effect on performance in the first year of the BA degree caused by the major taken at high school. From the perspective of the onto-semiotic model (Godino *et al.*, 2007; Godino *et al.*, 2019), the global problem is made up of different dimensions: epistemic, cognitive, interactional, ecological, affective and mediational. The epistemic dimension, which is not the object of study in the present work, is dealt with in other studies such as that of Contreras de la Fuente (2001), focused on the epistemological obstacles that arise in mathematics subjects in the initial years of university. In contrast, this research focuses on the cognitive dimension, although the other facets may also have some influence. The aim is to contribute to a better understanding of teaching and learning processes through the identification of the factors that may explain the differences in student performance and facilitate the implementation of measures to reduce academic failure.

Two samples were taken from universities that are very different in their ecological dimension, especially in terms of the socio-demographic profile of the students, and have a larger size than used in previous research. The results indicate that the value of the coefficient relating the high school major to the average mark in Business Mathematics I and II is practically identical at both universities: 0.54/0.56 (regression model and neural network respectively) for the Universidad Complutense de Madrid, and 0.56/0.52 for the Universidad Pontificia Comillas. Taking into consideration that we have worked with two different methodological approaches and two different universities, this consistency suggests that the results are robust. The conclusion is that the major taken at high school has a considerable impact on performance in mathematics in the first year of the BA degree, with students in the science major performing better.

From the point of view of the educational pathway (interactional dimension), it is surprising that students who have taken the recommended major to study Business Administration (social sciences) are at a disadvantage compared with those who opted for a non-recommended major (sciences). This effect extends to other related subjects, such as Business Statistics, a subject in which there is surprisingly no difference

between the two groups, despite the fact that the statistical workload is greater in the social sciences major. Given that the risk of academic failure is precisely concentrated in these subjects, these results invite reflection on the design of high school curricula. However, it should not be taken for granted that university programmes are not in need of revision (Ellerton and Clements, 1998).

In the cognitive dimension, a possible explanation of the result, supported by Skemp's (1976) approach, is that mathematics in the social science major has an instrumental focus, while mathematics in the science major achieves a greater relational understanding. Using the terminology of constructivist theory, relational understanding facilitates learning outcomes already achieved to scaffold new learning objectives. In this way, the zone of proximal development (Vygotsky, 1996) is expanded, so that they can be expected to understand new concepts more efficiently and achieve better results. It can be understood that an indicator of relational understanding is that there is a correlation between the student's results in one subject and those on which it is based (Gimenez *et al.*, 2021). Therefore, in order to evaluate at least indirectly the plausibility of this explanation, in Annex II we have performed an exercise by estimating the regression models of the mark in Mathematics II against the mark in Mathematics I, and incorporating the mark in EvAU and gender as control variables. For students from the science major, the coefficient associated with the mark in Mathematics I is 0.61 and 0.56 (Universidad Complutense de Madrid and Universidad Pontificia Comillas respectively). In the case of students taking social sciences, this coefficient is lower, 0.43 and 0.44 respectively. This seems to confirm the hypothesis that students taking the science major have a better relational understanding in mathematics, and perhaps this is one of the factors that explains their higher performance in the first year of the degree.

The difficulties in learning mathematics in the first year of university may also stem from the affective dimension and its interaction with the cognitive facet (Gómez-Chacón, 2009). A second possible explanation for the results of this work, compatible with the first, is that the motivations and skills that led students to choose their high school major are related to their skills and attitudes towards abstract thinking and, consequently, to their performance in mathematics subjects. The results of Corrales Serrano (2020) support this explanation to some extent, since the personal perception of abilities and aptitudes appear as the third most important criterion

Received: 15-10-2022 Accepted: 30-05-2023

for students to select their high school major (after two internal criteria on the valuation of the profession which it will give them access to).

From the point of view of teaching practice, the results of this work have two implications. Firstly, and with regard to the approach and content of mathematics during high school, it would be necessary to reflect on the social sciences major. It is interesting that students taking this major, which in general is the one recommended for BA studies, obtain worse results than those who took a major recommended for other degrees. Therefore, it seems appropriate to review the focus and content of mathematics in this major, seeking greater alignment with the requirements and needs of the related degrees. Secondly, we can reflect on teaching practice in the first year of university. As indicated, the theory of didactic suitability presented in Godino *et al*. (2021) facilitates the identification of criteria that contribute to good teaching practices. In the cognitive dimension, it highlights that the instructional process supports students according to their individual differences in prior knowledge and learning styles as a criterion of suitability. Therefore, better knowledge of the factors that generate differences among students facilitates compliance with this criterion. As has been shown in this work, the major taken at high school is one of these factors, the impact of which is considerable. Consequently, it seems necessary to rethink the didactic strategies in mathematics used during the first year of university. Perhaps it would be appropriate to consider the characteristics and learning style of students taking the social sciences major, adapting the teaching methodology to them.

Limitations and future lines of research

This study has two main limitations. Firstly, the study was carried out on a single degree programme. Given that the main objective of this research was to perform a comparative analysis between two universities, it is difficult to find degrees that are sufficiently similar for such a comparison to be suitably made. Even in the case studied in this paper, the analysis of the subject Business Statistics presents the problem that it is taught in different years (first year at the Universidad Complutense de Madrid and second year at the Universidad Pontificia Comillas), which prevents direct comparison, as mentioned above. In this regard, as a

future line of research, it is proposed to extend the analysis to other degrees, such as psychology or social work, degrees that are currently offered at the two universities analyzed and that include quantitative subjects in their curricula.

A second limitation is the number of universities compared. It would be desirable to extend the study to more universities, particularly in other Spanish provinces, in order to achieve a broader comparison. In this regard, it would be particularly relevant to replicate the work, using more recent samples, in those faculties in which discrepant results with those obtained in this work have been obtained, as is the case of the University of Las Palmas de Gran Canaria (Dávila *et al.*, 2015).

Finally, it is proposed as a future line of research to study the evolution throughout the degree in other quantitative subjects, such as Econometrics and Financial Mathematics. It would be of interest to check whether the advantage that students from the science major have in the mathematical subjects in the first year is maintained or disappears in mathematical subjects in later years. The results obtained for the case of Business Statistics seem to indicate that this advantage tends to disappear: at the Universidad Complutense de Madrid, where it is taken in the first year, there are differences at a significance level of 5%, while at the Universidad Pontificia Comillas, where it is taken in the second year, there are none. However, precisely because it is not a homogeneous comparison, this conclusion is merely indicative. It would be necessary to carry out a more in-depth study that would allow a realistic comparison both within each center and between different centers.

Bibliographical references

Arroyo-Barrigüete, J. L., Tirado, G., Mahíllo Fernández, I., & Ramirez, P. (2020a). Predictors of performance in Business Administration degrees: the effect of the high-school specialty. *Revista de educación, 390*, 129-154. https://doi.org/10.4438/1988-592X-RE-2020-390-468

Arroyo-Barrigüete, J. L., Carabias López, S., Curto González, T. & Borrás Palá, F. (2020b). Matemáticas en el doble grado ADE-Derecho: un análisis cuantitativo de las estrategias de estudio. Bordón: *Revista de pedagogía, 72*(4), 27-42. https://doi.org/10.13042/Bordon.2020.80306

Received: 15-10-2022 Accepted: 30-05-2023

Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. (2014). NbClust: An R Package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*(6), 1-36, https://doi.org/10.18637/jss.v061.i06

Contreras de la Fuente, Á. (2001). La enseñanza del análisis matemático en el bachillerato y primer curso de universidad: una perspectiva desde la teoría de los obstáculos epistemológicos y los actos de comprensión. En *IV Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 71-85).

Corrales Serrano, M. (2020). *Estudio de las motivaciones internas y externas para la elección de modalidad del alumnado de Bachillerato. Incidencia en la didáctica de las Ciencias Sociales*. [Tesis Doctoral]. Universidad de Extremadura. https://dehesa.unex.es/handle/10662/11076

Dávila, N., García-Artiles, M., Pérez-Sánchez, J. M., & Gómez-Déniz, E. (2015). Un modelo de regresión logística asimétrico que puede explicar la probabilidad de éxito en el rendimiento académico. *Revista de Investigación Educativa, 33*(1), 27-45. https://doi.org/10.6018/rie.33.1.178481

De Guzmán, M., Hodgson, B. R., Robert, A., & Villani, V. (1998, August). Difficulties in the passage from secondary to tertiary education. In *Proceedings of the International Congress of Mathematicians* (Vol. 3, pp. 747-762). Berlin: Documenta Mathematica.

Ellerton, N. F., & Clements, M. A. (1998). Transforming the international mathematics education research agenda. En *Mathematics education as a research domain: A search for identity* (pp. 153-175). Springer.

Ernest, P. (2016). An Overview of the Philosophy of Mathematics Education. En *The philosophy of mathematics education* (3-8). Springer Open. https://doi.org/10.1007/978-3-319-40569-8_1

Fox J., &Weisberg, S. (2019). *An {R} Companion to Applied Regression*, Third Edition. Sage.

Gaure, S. (2013). OLS with multiple high dimensional category variables. *Computational Statistics & Data Analysis, 66*, 8-18. http://dx.doi.org/10.1016/j.csda.2013.03.024

Giménez, M. J., Borrás, F., & Carabias, S. (2021). Sesgo matemático negativo en los alumnos del doble grado en ADE y Derecho: Un análisis cuantitativo. En *Universidad, innovación e investigación ante el horizonte 2030* (p. 112). Egregius.

Godino, J. D., & Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en didactique des Mathématiques, 14*(3), 325-355.

Godino, J. D., Batanero, C., Burgos, M., & Gea, M. M. (2021). Una perspectiva ontosemiótica de los problemas y métodos de investigación en educación matemática. *Revemop, 3*, 1-30. https://doi.org/10.33532/revemop.e202107

Godino, J.D., Batanero, C. & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM Mathematics Education 39*, 127-135. https://doi.org/10.1007/s11858-006-0004-1

Godino, J. D., Batanero, C., & Font, V. (2019). The onto-semiotic approach. *For the learning of mathematics, 39*(1), 38-43.

Gómez-Chacón, I. M. (2009). Actitudes matemáticas: propuestas para la transición del bachillerato a la universidad. *Educación matemática, 21*(3), 5-32.

Gonzalez Veiga, M. C., Castellanos Val, L., Gonzalez De Sela Aldaz, M. A., Manzano Pérez, & I. M. (1999). Influencia del currículo previo sobre el rendimiento en matemáticas empresariales: un estudio para la universidad de Oviedo. *Revista de Educación, 319*, 223-238.

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software, 40*(3), 1-25. https://doi.org/10.18637/jss.v040.i03

Gueudet, G. (2008). Investigating the secondary–tertiary transition. *Educational studies in mathematics, 67*(3), 237-254. https://doi.org/10.1007/s10649-007-9100-6

Jaworski, B. (1998). Mathematics teacher research: process, practice, and the development of teaching. *Journal of Mathematics Teacher Education, 1*, 3-31. https://doi.org/10.1023/A:1009903013682

Kuhn, M. (2020). *caret: Classification and Regression Training*. R package version 6.0-86. https://CRAN.R-project.org/package=caret

Lang, D. T., & CRAN team (2019). *Package 'RCurl'*. URL http://www.omegahat.net/RCurl

Martínez de Ibarreta, C., Rua, A., Redondo, R., Fabra, M. E., Nuñez, A., & Martín, M. J. (2010). Influencia del nivel educativo de los padres en el rendimiento académico de los estudiantes de la ADE. Un enfoque de género. En M. J. Mancebón Torrubia, D. Pérez Ximénez de Embún, J. M. Gómez Sancho y G. Giménez Esteban (Coords.), *Investigaciones*

de Economía de la Educación Número 5 (pp. 1273-1296). AEDE, Asociación de Economía de la Educación.

Pizarroso, J., Portela, J., & Muñoz, A. (2022). NeuralSens: sensitivity analysis of neural networks. *Journal of Statistical Software, 102*(7), 1-36. https://doi.org/10.18637/jss.v102.i07

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL: http://www.R-project.org/

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching,* 77, 20-26.

Vygotsky, L. S. (1996). *El desarrollo de los procesos psicológicos superiores*. Cole, M., John-Steiner, V., Scribner, S., Souberman, E. (eds); Crítica. (Original publicado en 1978).

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., & Venables, B. (2020). *gplots: various R programming tools for plotting data*. R package version 3.1.1. https://CRAN.R-project.org/package=gplots

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wickham, H. François, R., Henry, L., & Müller, K. (2022). *dplyr: a grammar of data manipulation*. R package version 1.0.8. https://CRAN.R-project.org/package=dplyr

Zeileis, A., & Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. *R News, 2*(3), 7-10. https://CRAN.R-project.org/doc/Rnews/

Contact address: José Luis Arroyo-Barriguete. Universidad Pontificia Comillas. Facultad de Ciencias Económicas y Empresariales, departamento de Métodos Cuantitativos. Calle de Alberto Aguilera, 23, 28015, Madrid, Spain. E-mail: jlarroyo@comillas.edu

Annex 1: cluster analysis at Universidad Complutense de Madrid

In order to verify the importance of Mathematics I and II in the first-year subjects, a sample of 3,279 students of the BA degree at the Universidad Complutense de Madrid was taken as a starting point. We have considered the cohorts from the academic year 2009/2010 (the year in which the current study plan came into force) until 2021/2022. A large number of records do not have information on the high school major, whether sciences or social sciences, which is why they could not be used in the main analysis. However, in all of them, there is information on the mark obtained by each student in the first exam sitting of the 10 subjects comprising the Basic Training[9]. The sample of 3,279 students is summarized in Table VI. Therefore, a cluster analysis has been performed considering these marks as variables.

TABLE VI. Sample used for the cluster analysis (3,279 students on the BA degree at the Universidad Complutense de Madrid): average mark in each subject

Financial Accounting I	**Fundamentals of Business Administration**	**Fundamentals of Financial Administration**	**Company Law**	**Introduction to Economics**
4.33	5.21	3.70	5.29	4.34
Microeconomics	**Business Mathematics I**	**Business Mathematics II**	**Business Statistics I**	**Economic History**
3.95	2.83	3.24	3.70	4.53

The Euclidean distance and Ward's linkage method were chosen. The grouping by subjects (vertical interpretation of Figure III), points to three different clusters, one of which is comprised of Business Mathematics I and II. The first cluster, made up of Company Law, Economic History, Introduction to Economics, and Fundamentals of Business Administration, seems to be made up of subjects of comprehension with a certain rote learning nature, whose quantitative weight is reduced or null. The second cluster is made up of Business Statistics I, Microeconomics, Fundamentals of Financial Administration, and Financial Accounting I, subjects of a quantitative nature, although not purely mathematical. The last cluster includes Business Mathematics I and II. Analysis of the average marks in each of

[9] Non-attendance at the exam sitting equates to a mark of 0 for the subject.

Received: 15-10-2022 Accepted: 30-05-2023

these groups confirms considerable differences, as shown in Table VII: the average mark in Business Mathematics I and II is not only significantly lower than in other clusters, but the distribution is markedly different.

FIGURE III. Heatmap for the 10 Basic-Training subjects at Universidad Complutense de Madrid

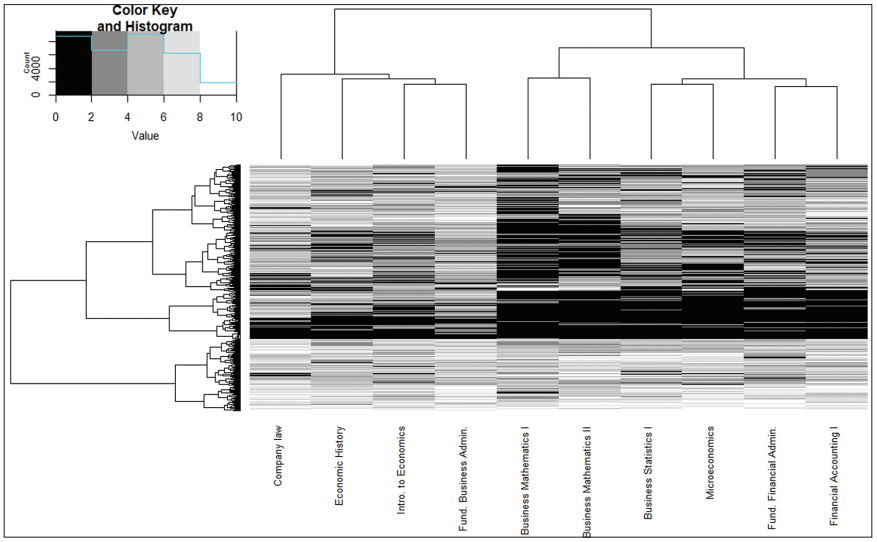
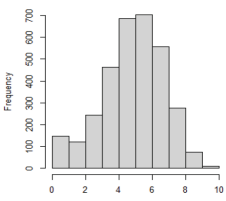
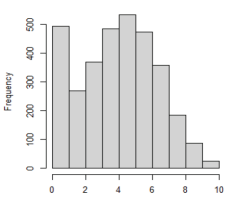
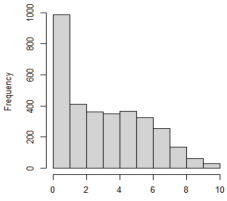


TABLE VII. Main statistics for the average mark in each cluster identified at Universidad Complutense de Madrid

	Cluster 1: comprehension subjects	Cluster 2: quantitative subjects	Cluster 3: Business Mathematics I and II
First quartile	3.75	2.20	0.65
Median	4.98	4.05	2.75
Mean	4.84	3.92	3.04
Third quartile	6.18	5.63	5.00
Histograms			

Annex 2: relationship of performance in Business Mathematics I and II

Table VIII shows the results of the regression model using the mark in Mathematics II as the dependent variable and the result in Mathematics I as the independent variable, including gender and the EvAU mark as control variables. In this case, we have considered the sample of students from the Universidad Complutense de Madrid, split into two groups: students from the social sciences major and students who took the science major. Table IX is identical except considering the sample of students from the Universidad Pontificia Comillas. As with the previous models, the effect of the high school major is remarkably similar in both samples. It is observed that, in the case of students from the science major, the effect of the mark in Mathematics I is higher than for students from the social sciences. This seems to confirm, at least indirectly, the hypothesis that students from the science major have a better relational understanding of mathematics.

TABLE VIII. Regression model for the variable Mathematics II (Universidad Complutense de Madrid), distinguishing by high school major

	High school major: sciences (R^2 = 0.45)			**High school major: social sciences (R^2 = 0.27)**		
	Coef.	**sd**	**p-value**	**Coef.**	**sd**	**p-value**
Constant	-0.15	0.07	3.5 E-2	-0.08	0.04	8.0 E-2
Gender (Female=1)	0.45	0.13	4.9 E-4	0.17	0.07	1.2 E-2
EvAU	0.00	0.06	9.6 E-1	0.13	0.03	2.8 E-4
Mark Mathematics I	0.61	0.06	< 2 E-16	0.44	0.03	< 2 E-16

TABLE IX. Regression model for the variable Mathematics II (Universidad Pontificia Comillas), distinguishing by high school major

	High school major: sciences (R^2 = 0.39)			**High school major: social sciences (R^2 = 0.36)**		
	Coef.	**sd**	**p-value**	**Coef.**	**sd**	**p-value**
Constant	-0.01	0.07	8.6 E-1	-0.07	0.05	1.7 E-1
Gender (Female=1)	0.03	0.10	7.9 E-1	0.14	0.07	5.8 E-2
EvAU	0.13	0.05	1.5 E-2	0.24	0.04	1.8 E-8
Mark Mathematics I	0.56	0.07	1.4 E.15	0.43	0.04	< 2 E-16

Received: 15-10-2022 Accepted: 30-05-2023

Community engagement: A bibliometric analysis in the university context

Community engagement: Un análisis bibliométrico en el contexto universitario

https://doi.org/10.4438/1988-592X-RE-2023-402-598

Sebastián Araya-Pizarro
http://orcid.org/0000-0002-5857-8441
Universidad de La Serena

Nando Verelst
https://orcid.org/0000-0002-0957-767X
Universidad de La Serena

Abstract

Modern universities have the task of transferring skills, knowledge and values, that serves the students to engage with societal and environment challenges. This research aims to examine the research landscape in the field of university-community engagement. M***ethodology:*** A longitudinal descriptive bibliometric examination of the scientific literature published during the period 2017-2021, was applied. Data were obtained through the search in the Scopus and Web of Science databases, using the PRISMA declaration for systematic reviews. Bibliometric indicators were calculated to describe the production, dispersion, distribution and growth of the articles. Through bibliometric mapping, the co-occurrence networks of the most used terms and citations were explored. A thematic analysis was also performed using the categorization scheme of Callon et al. (1991). To analyze the data, the Bibliometrix R package and the Biblioshiny application were used, and the VOSviewer software were used to visualize the outputs of the network analysis. ***Results:*** There were 404 articles, which received 431 citations, mostly from the United States of America (277), with an average of 8 citations per document. The annual growth of the literature was 29%, with an exponential adjustment trend (R2= .93). Most of the authors published only one article (93%), mostly in the areas of education and health.

Received: 28-08-2022 Accepted: 11-05-2023

Eight thematic clusters were found, distributed into motors, basics, niche and emerging topics. ***Conclusions:*** The study confirms a growing scientific production on community engagement, which is concentrated in few journals and numerous authors. Various topics are denoted (such as community engagement and development, service learning, and community-based participatory research) that have relevant implications for universities, higher education institutions and policymakers.

Keywords: community engagement, bibliometric analysis, university, Scopus, Web of Science.

Resumen

Las universidades modernas tienen la tarea de transferir habilidades, conocimientos y valores que permitan al estudiantado lograr un mayor compromiso con los desafíos de la sociedad y el medio ambiente. Esta investigación tiene como objetivo analizar el panorama investigativo del vínculo universidad-comunidad (*community engagement*). ***Metodología:*** Se aplicó un examen bibliométrico descriptivo longitudinal de la literatura científica publicada durante el periodo 2005-2021. Los datos fueron obtenidos mediante la búsqueda en las bases de datos *Scopus* y *Web of Science*, empleando la declaración PRISMA para revisiones sistemáticas. Se calcularon indicadores bibliométricos para describir la producción, dispersión, distribución y crecimiento de los artículos. A través del mapeo bibliométrico, se exploraron las redes de coocurrencia de los términos claves más utilizados. También se realizó un examen temático mediante el esquema de categorización de Callon et al. (1991). Para analizar los datos se utilizó el paquete *Bibliometrix* R y la aplicación *Biblioshiny* y, para la visualización de las salidas del análisis de red se usó el programa *VOSviewer*. ***Resultados:*** Hubo 404 artículos, que fueron publicados en 262 revistas, provenientes mayormente de Estados Unidos (277), con un promedio de 8 citas por documento. El crecimiento anual de la literatura fue de 29%, con una tendencia de ajuste exponencial (R^2= .93). La mayoría de los autores publicó solo un artículo (93%) en las áreas de educación y salud. Se hallaron 8 clústeres temáticos distribuidos en tópicos motores, básicos, de nicho y emergentes. ***Conclusiones:*** El estudio confirma una producción científica creciente sobre *community engagement*, que se concentra en pocas revistas y numerosos autores. Se denotan diversos tópicos (como el compromiso y desarrollo comunitario, el aprendizaje-servicio y la investigación participativa basada en la comunidad) que tienen implicaciones relevantes para universidades, instituciones de educación superior y responsables de políticas públicas.

Palabras clave: community engagement, análisis bibliométrico, universidad, Scopus, Web of Science.

Introduction

There is a growing demand for the review and reform of educational curricula in higher education, with the purpose of moving towards a more globalized and integral formation, according with the challenges inherent to the sustainable development of humanity (Chen, 2021; Zapp & Lerch, 2020). Thus, higher education institutions (HEIs) have the task of transferring skills, knowledge, and values that allow students and future professionals to achieve a significant link with social and environmental needs (Robertson, 2021).

In the academic context, community engagement (CE) seeks to establish a sustainable, holistic, and reciprocal relationship between academia and the community (Sheila et al., 2021). This relationship is materialized through the execution of relevant research projects for community stakeholders (Machimana et al., 2021) and active learning activities that encourage student engagement with diverse communities (Bidandi et al., 2021). In this regard, universities have the potential to contribute to the development of social (relational) capital through actions of creation and knowledge transfer in the different disciplinary fields in which they operate (Breznitz & Feldman, 2012; Johnston & Lane, 2018). Specially, considering that the notion of "engaged university" has been distinguished as a plausible model for higher education (Dougnac, 2016; Johnston et al., 2021).

Despite the transversal recognition of the importance of CE, its development in the university activity is still incipient (Bidandi et al., 2021) and its relevance, as a research topic, is recent. This provides the opportunity to build a broad and detailed overview of academic research on CE that synthesizes the state of the art, trends over time, key topics, and future trends in this area.

Bibliometric analysis, on the other hand, is a method for studying large volumes of scientific data based on semantic units such as keywords, and production units such as countries, institutions, journals, and authors (Aria & Cuccurullo, 2017; Donthu et al., 2021). In addition, it has the potential to define relevant trends and themes (both primary and emerging) through the application of scientific mapping techniques, providing a general view of the field of study (Callon et al., 1991).

This article aims to analyze the research landscape on the study of community engagement in universities, based on the bibliometric review of the two leading bibliographic databases in the world: Web of Science

Received: 28-08-2022 Accepted: 11-05-2023

(WoS) and Scopus (Zhu & Liu, 2020). The analysis horizon includes the last 17 years (2005-2021), a period in which a continuous production of scientific research on the subject is observed. In this way, it is expected to contribute to the understanding of the university-community engagement phenomenon with bibliometric evidence; a prominent topic, but characterized by heterogeneity in its conceptualization and territorial theoretical development (Compagnucci & Spigarelli, 2020; Dougnac, 2016).

Literature review

Community engagement is a collaborative work process with a group or groups of people to meet their needs or improve their well-being (Levin et al., 2021). According to Johnston and Lane (2018), this process is intimately linked to the concept of relational capital, a fundamental resource for the functioning of a democratic society.

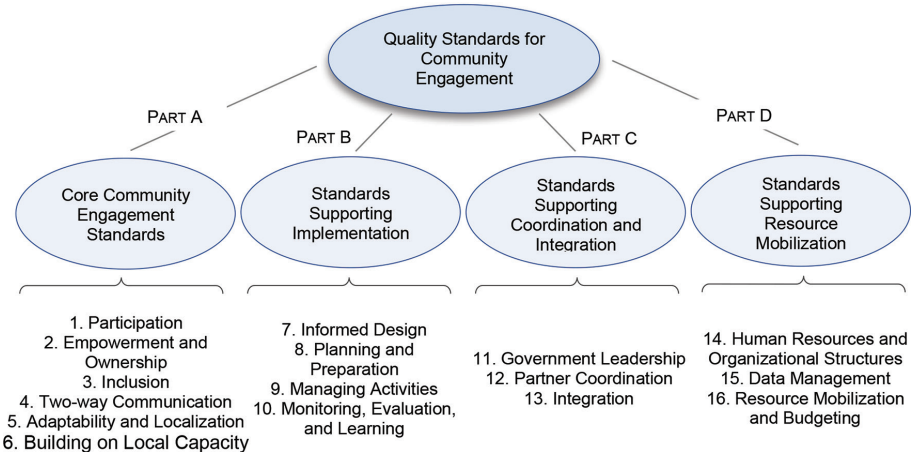
In this line, universities CE is a topic that is gaining global relevance (Groark & McCall, 2018). It is understood as a process that brings together actors from civic associations and non-profit organizations, public institutions, private organizations, and universities, to form sustainable relationships aimed at improving the collective well-being of community stakeholders through collaborative actions and programs. Examples of such actions include project-based learning, service-learning, and applied research on socially relevant topics (Unger & Polt, 2017).

For this reason, it is argued that the systematic application of CE actions by universities strengthens their social responsibility since, by addressing the local and regional concerns of the territory, institutions manifest their commitment to the sustainable development of a country. Likewise, CE contributes to improving students' academic, social and employability skills, through active learning strategies (Bhatnagar et al., 2020; Mbah et al., 2021).

Although there is no global consensus on what are the basic dimensions that determine the university-community link, it is important to note that the United Nations Children's Fund (UNICEF) established 16 minimum quality indicators that serve as benchmarks to guide the development of CE programs and actions in humanitarian contexts (Figure I). Other reference frameworks for CE have also been proposed, mainly in the health area, which share elements with the model proposed by UNICEF (Lavery et al., 2010; McNaughton & Duong, 2014; UNICEF, 2020).

Received: 28-08-2022 Accepted: 11-05-2023

FIGURE I. UNICEF minimum quality standards for community engagement

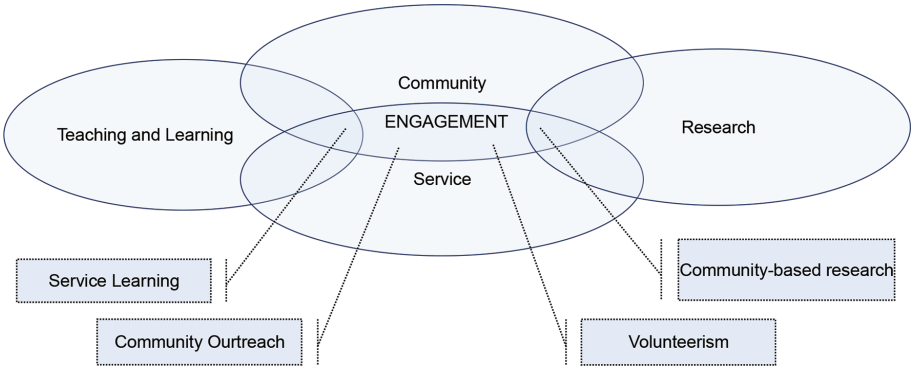


Source: UNICEF (2020).

Regarding higher education and university management, different studies have systematized and analyzed topics such as university-industry collaboration and entrepreneurial universities (Forliano et al., 2021; Skute et al., 2019; Zavale & Langa, 2018). The article by Compagnucci and Spigarelli (2020) stands out by examining the concept of the third university mission (contribution to the community and territories) and concluding that the university-community engagement must necessarily be interpreted as a bidirectional and dialogic process with the communities. Additionally, the authors highlight the high heterogeneity existing in the concepts involved, the interpretation of their function, the way in which they can be addressed, and the impacts they can generate.

On the other hand, the theoretical proposal of Bender (2008) recognizes three general roles of higher education institutions (teaching and learning, research, and community service) that are connected through interaction interfaces and incorporate the community as a preponderant factor, conforming a model that defines general CE strategies consistent with the functions of HEIs (Figure II).

FIGURE II. Theoretical model of CE in higher education



Source: Bender (2008).

In this sense, it is relevant to mention that the concept of Community Engagement involves the processes of linking organizations with communities, community participation in activities of general interest, and institutional commitment to the sustainable development of the community and territory (Richardson et al., 2021).

Moreover, some articles have explored the CE research field through bibliometric analysis, although mostly limited to the health area. Overall, all of these studies highlight the importance of joint working between public institutions, leaders, and community members towards a goal of common interest. For example, Yuan et al. (2021) examined, through bibliometric mapping techniques on the Web of Science, the literature on CE in public health. Gilmore et al. (2020) and Osborne et al. (2021) reviewed the field of CE in relation to strategies for preventing infectious diseases. Among the databases consulted by these authors are the Web of Science (Osborne et al., 2021) and Scopus (Gilmore et al., 2020). On the other hand, Ryan et al. (2020) explored, in Web of Science, Scopus, and other directories, the linkage of institutions with the community for natural disaster preparedness. In this way, this study provides updated and significant information on CE from a different and underexplored area, university education, based on the review of Web of Science and Scopus repositories, platforms widely highlighted for their visibility, scope, and international recognition.

Received: 28-08-2022 Accepted: 11-05-2023

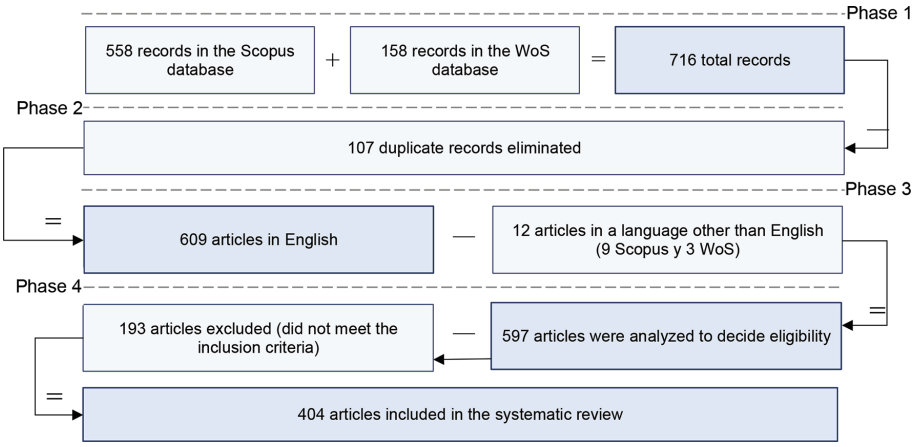
Method

The methodological process was based on the standardized workflow proposed by Zupic and Čater (2015), which establishes a procedure of five sequential phases: study design, data collection, data analysis, data visualization, and data interpretation.

The study design is based on a quantitative approach, descriptive type and longitudinal scope. To measure the current state of the field of study on CE and its development over time, a bibliometric analysis and network mapping of the literature published during the period 2005-2021 was applied.

Data was collected by searching the two most recognized bibliographic databases worldwide, WoS and Scopus (Zhu & Liu, 2020), using the PRISMA statement for systematic reviews (Moher et al., 2009). In phase 1 involved the search for articles took place; in phase 2 duplicate publications were removed; and in phase 3, articles that did not meet the inclusion criteria were discarded. Article eligibility parameters were decided in phase 4, focusing on the premise that they were documents written in English that dealt with the university CE. The review in WoS yielded 158 documents, while Scopus yielded 558. The search was conducted on August 5th, 2022. Both datasets were merged using a script written in R language, leaving the final database composed of 404 documents (Figure III).

FIGURE III. Study selection flow according to the PRISMA Declaration



Source: Compiled by author.

The search and selection strategy were carried out through the application of Boolean terms ("university" OR "community engagement") AND ("community engagement" OR "community participation" OR "community involvement" OR "community consultation"), using titles and keywords. This is in line with the work developed by Yuan et al. (2021). The option to incorporate only studies in English is based on the difficulty of articulating an accurate translation of the Anglo-Saxon concept engagement (Dougnac, 2016).

The inclusion criteria used were: 1) empirical studies published in scientific journals; 2) scientific articles that were in their final publication stage; 3) articles published from January 1, 2005 to December 31, 2021 (inclusive); and 4) studies focused on investigating university-community engagement.

For the analysis of collected data, a compilation of bibliometric techniques and network mapping was applied (Cobo et al., 2011; Muñoz-Leiva et al., 2012). Indicators describing the production, dispersion, distribution, and evolution of articles were calculated. Furthermore, the growth of scientific literature (Price's law), the concentration of authors (Lotka's law), and the Bradford model were calculated.

Moreover, through bibliometric mapping, the co-occurrence networks of the most used terms were explored, and an analysis of keyword trends and a thematic examination were carried out using the categorization scheme of Callon et al. (1991). This algorithm, by estimating the parameters of centrality (degree of interaction of the network with other networks) and density (strength of internal associations of a group), categorizes research themes into four dimensions: motor themes (high centrality and density), basic themes (high centrality and low density), niche themes (low centrality and high density), and emerging or declining themes (low centrality and density).

Finally, it should be noted that Bibliometrix package and Biblioshiny app, both developed by Aria and Cuccurullo (2017), were used to analyze the data. While for the visualization of the bibliometric networks maps the VOSviewer program was used.

Results

Sample characterization

In total, 404 documents from 262 journals were retrieved. The annual growth rate was 28.9%, with an average of 8 citations per article. A total

Received: 28-08-2022 Accepted: 11-05-2023

of 1,582 automatically-generated keywords (Keyword Plus) and 1,158 author keywords were identified. Additionally, 1,424 authors contributed to the field of CE, with 69 of them being single-author document authors. The average number of co-authors per document was 3.8, and the international co-authorship rate was 5% (Table I).

TABLE I. Sample characterization

Dimension	Descriptor	Result
Scientific production	Timespan	2005:2021
	Sources (Journals, Books, etc.)	262
	Annual Growth Rate %	28.89
Documents	Total documents	404
	Document Average Age (years)	5.03
	Average citations per document	8.015
	Keywords Plus (ID))	1,582
	Author's Keywords (DE)	1,158
Authors	Authors	1,424
	Authors of single-authored documents	69
Collaboration	Single-authored documents	72
	Co-Authors per documents	3.83
	International co-authorships (%)	4.95

Source: Compiled by author.

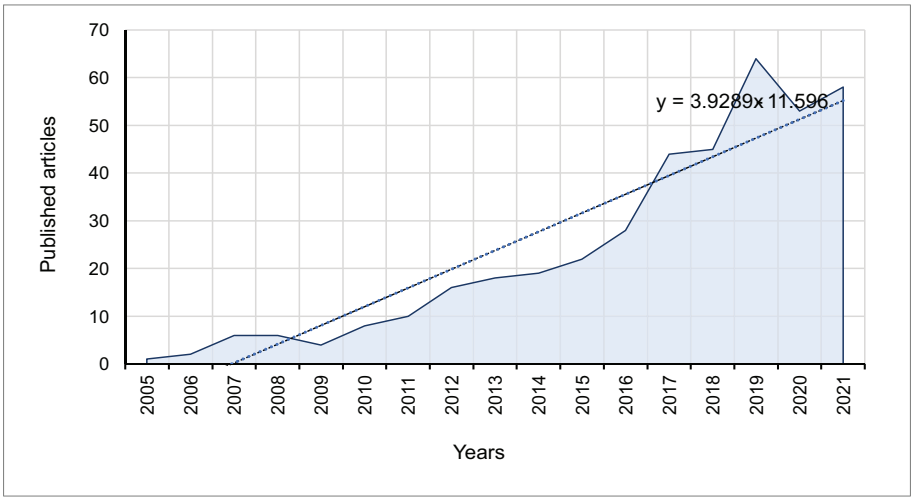
Temporal evolution and size of literature

The scientific production during the period 2005-2021 showed a positive trend, with an average annual growth of 29% (Figure IV).

The linear fit of the time series revealed an increase of four articles per year (β= 3.93, R^2= .89). Moreover, it is possible to confirm that the temporal evolution of scientific production mainly adheres to an exponential model (β= 0.23, R^2= .93), fulfilling the postulates of Price's Law of exponential growth of literature (Table II).

Received: 28-08-2022 Accepted: 11-05-2023

FIGURE IV. Evolution of scientific production on CE, period 2005-2021



Source: Compiled by author.

TABLE II. Linear and exponential estimation of scientific production growth

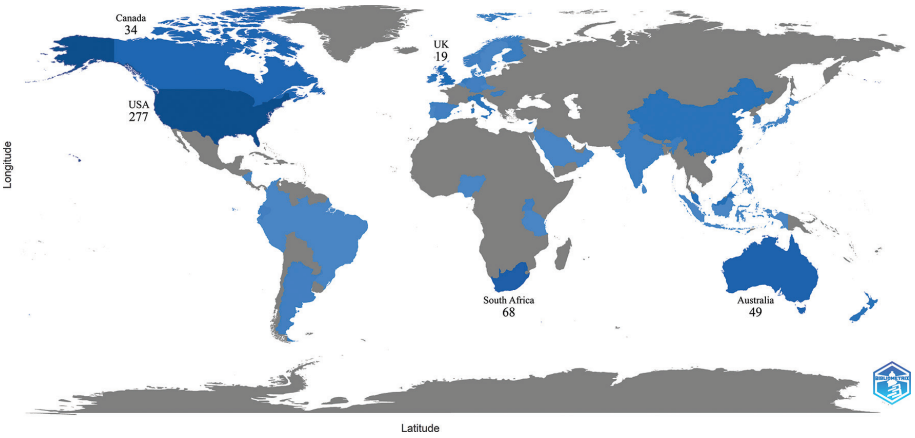
Equation	R^2	F	Significance	Constant	β
Linear	.888	118.53	.000***	-11.60	3.929
Exponential	.925	184.37	.000***	1.73	0.232

Note: y = Number of publications; x = Time periods (1-17), ***p< .001.
Source: Compiled by author.

Analysis of the geographic distribution

Figure V shows that the countries with the highest scientific production were, in decreasing order, the United States of America (277 articles), South Africa (68), Australia (49), Canada (34) and the United Kingdom (19). These results confirm the undisputed leadership of the United States, which would be explained by its territorial extension, which hosts important institutions, centers, and research groups recognized worldwide, as well as by the language in which they publish (English).

FIGURE V. Geographic distribution of scientific production



The map highlights in blue tones, from darkest to lightest, the countries with the highest to lowest scientific production on CE. Source: Compiled by author.

Institutional analysis

In the ranking of the 10 most productive institutions, the dominance of American (6) and South African (3) universities stands out, with a total of 57 and 20 publications, respectively. The most relevant institution was Johns Hopkins University School of Medicine (14), followed by the University of California (13) and the University of North Carolina (9) (Table III).

TABLE III. Top ten most productive institutions

Ranking	Institution	Country	Publications
1	Johns Hopkins University School of Medicine	USA	14
2	University of California	USA	13
3	University of North Carolina	USA	9
4	University of Pretoria	South Africa	8
5	University of British Columbia	Canada	8

(*Continued*)

Received: 28-08-2022 Accepted: 11-05-2023

TABLE III. Top ten most productive institutions *(Continued)*

Ranking	Institution	Country	Publications
6	University of Washington	USA	8
7	University of Wisconsin-Madison	USA	7
8	Durban University of Technology	South Africa	6
9	North West University	South Africa	6
10	Rowan University	USA	6

Source: Compiled by author.

Authors analysis

The main authors published a maximum of 4 articles each, demonstrates the low-density present in the authorship of scientific production on CE. The most productive researchers were Brown, Govender, Salzer, and Weerts (Table IV).

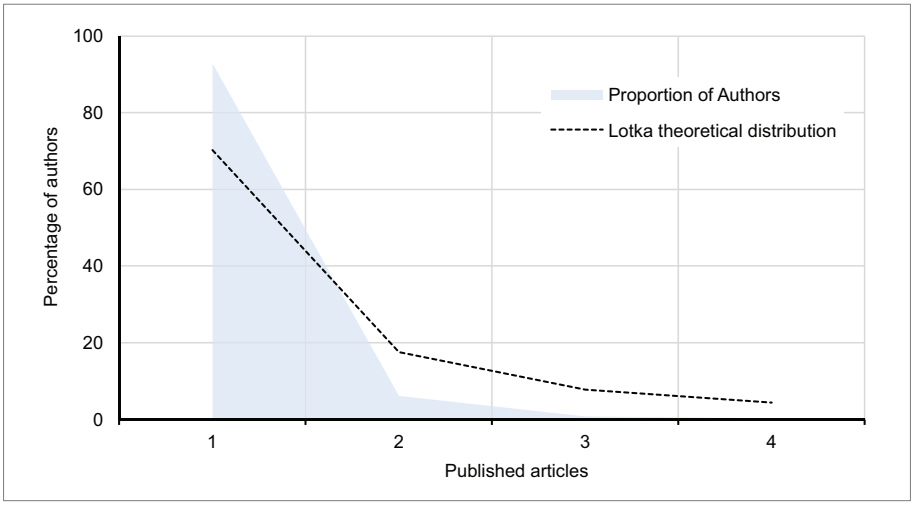
TABLE IV. Top ten most productive authors

Ranking	Author	Publications	Percentage	H-index	Total Citations	Citations per article
1	Brown, A.	4	1.0	3	43	10.8
2	Govender, I.	4	1.0	2	8	2.0
3	Salzer, M.	4	1.0	3	77	19.3
4	Weerts, D.	4	1.0	3	240	60.0
5	Brown, K.	3	0.7	3	30	10.0
6	Brusilovskiy, E.	3	0.7	2	72	24.0
7	Ebersohn, L.	3	0.7	2	13	4.3
8	Eloff, I.	3	0.7	1	6	2.0
9	Kruss, G.	3	0.7	2	32	10.7
10	Machimana, E.	3	0.7	1	8	2.7

Source: Compiled by author.

The examination of author concentration, through Lotka's Law (Lotka, 1926), confirmed the high dispersion present in the reviewed production. Specifically, it was determined that 92.8% of authors have published one document, 6.1% have published two articles, and only 1.1% have published more than two articles (Figure VI).

FIGURE VI. Lotka's Law



Source: Compiled by author.

Journals analysis

Table V shows that the journals that have published the most on CE belong to the fields of education and health. The most relevant sources were the Journal of Higher Education Outreach and Engagement (16), Progress in Community Health Partnerships: Research, Education, and Action (13), and Gerontology and Geriatrics Education (9).

Received: 28-08-2022 Accepted: 11-05-2023

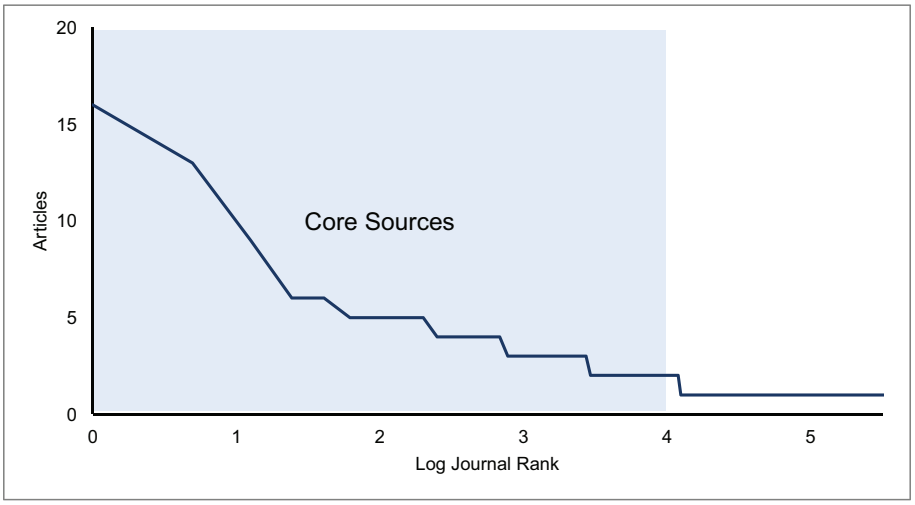
TABLE V. Top ten most productive journals

Ranking	Journal	Publications	Percentage	Total Citations	H-Index	Research domain
1	Journal of Higher Education Outreach and Engagement	16	4.0	39	3	Education
2	Progress in Community Health Partnerships: Research, Education, And Action	13	3.2	36	3	Education
3	Gerontology and Geriatrics Education	9	2.2	61	5	Education
4	Higher Education Research & Development	6	1.5	56	4	Education
5	International Journal of Environmental Research and Public Health	6	1.5	30	2	Public, Environmental and Occupational Health
6	American Journal of Community Psychology	5	1.2	59	3	Public, Environmental and Occupational Health
7	Educational Leadership	5	1.2	3	1	Education
8	Journal of Geography in Higher Education	5	1.2	32	4	Education
9	South African Journal of Higher Education	5	1.2	3	1	Education
10	South African Review of Sociology	5	1.2	48	3	Social Sciences

Note: Total citations and H-Index (Source Local Impact), Research domain by Scimago.
Source: Compiled by author.

The analysis of productivity according to publication sources showed a concentrated distribution. Specifically, the calculation of Bradford's Law (Bradford, 1985) revealed a high degree of correlation (R^2 = .97) that shows that most publications on CE are present in a small nucleus of journals (Figure VII).

FIGURE VII. Source grouping through Bradford's Law



Source: Compiled by author.

Publications citation analysis

The most cited article (125) was "Community Engagement and Boundary-Spanning Roles at Research Universities" by Weerts and Sandmann (2010). The same authors published the second most cited article, "Building a Two-Way Street: Challenges and Opportunities for Community Engagement at Research Universities" (Weerts & Sandmann, 2008). Both studies explore the concept of community engagement promoted by universities carrying out research actions. Studies by Winter et al. (2006), "University-community engagement in Australia", and Bender (2008), "Exploring conceptual models for community engagement at higher education institutions in South Africa", analyze CE practices and models in higher education institutions in Australia and Africa, respectively. On the other hand, Pommier et al. (2020) and Salzer et al. (2014) address methodological aspects for measuring of community engagement (Table VI).

TABLE VI. Top ten most cited articles

Ranking	Article title	Author (year)	Total Citations	Citations per year
1	Community Engagement and Boundary-Spanning Roles at Research Universities.	Weerts & Sandmann (2010)	125	9.6
2	Building a two-way street: Challenges and opportunities for community engagement at research universities.	Weerts & Sandmann (2008)	104	6.9
3	The Development and Validation of the Compassion Scale.	Pommier et al. (2020)	58	19.3
4	University-community engagement in Australia.	Winter et al. (2006)	57	3.4
5	Measuring community participation of adults with psychiatric disabilities: Reliability of two modes of data collection.	Salzer et al. (2014).	53	5.9
6	Exploring conceptual models for community engagement at higher education institutions in South Africa.	Bender (2008)	53	3.5
7	Outcomes of Australian Rural Clinical Schools: A Decade of Success Building the Rural Medical Workforce Through the Education and Training Continuum.	Greenhill et al. (2015)	52	6.5
8	The scholarship of university-community engagement: Interrogating Boyer's model.	Mtawa et al. (2016)	47	6.7
9	Town-gown relationships: Exploring university-community engagement from the perspective of community members.	Bruning et al. (2006)	46	2.7
10	Benefits and challenges of incorporating citizen science into university education.	Mitchell et al. (2017)	42	7.0

Source: Compiled by author.

Keyword analysis

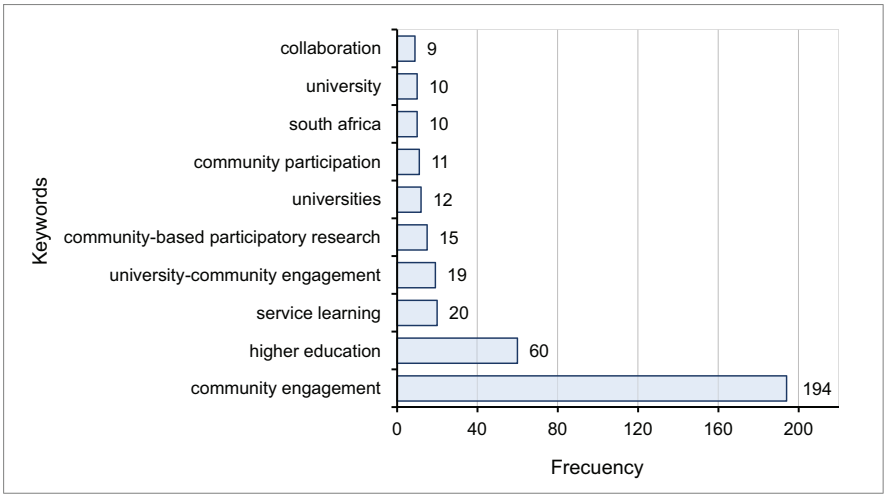
The main keywords used by the authors were: community engagement (194), higher education (60), service learning (20), university-community engagement (19), community-based participatory research (15), universities (12), and community participation (11) (Figure VIII).

Received: 28-08-2022 Accepted: 11-05-2023

The bibliographic coupling analysis, according to the consulted database, confirmed the predominance of the keywords community engagement, higher education, and service-learning, as observed in the density visualization maps in Figure IX.

Figure X shows the co-occurrence network of author keywords. In the bibliometric map, the size of the nodes is proportional to the frequency of occurrences and their relative weight, through which the terms were related to each other. The central clusters of the map indicate a high interrelation with the basic concepts that conform it, while those located at the ends indicate a low interrelation. Thus, five central nodes of co-words (community engagement, higher education, community-based participatory, audience and social justice) were registered in the network, whose examination (degree of centrality of the network) ratified the term CE as the main source for interaction and linkage with the other keywords of the reviewed articles. This topic recorded levels of intermediation and closeness of 936.3 and 0.019, respectively, much higher than other concepts such as higher education (54.1 and 0.012) or community-based participatory (31.6 and 0.011).

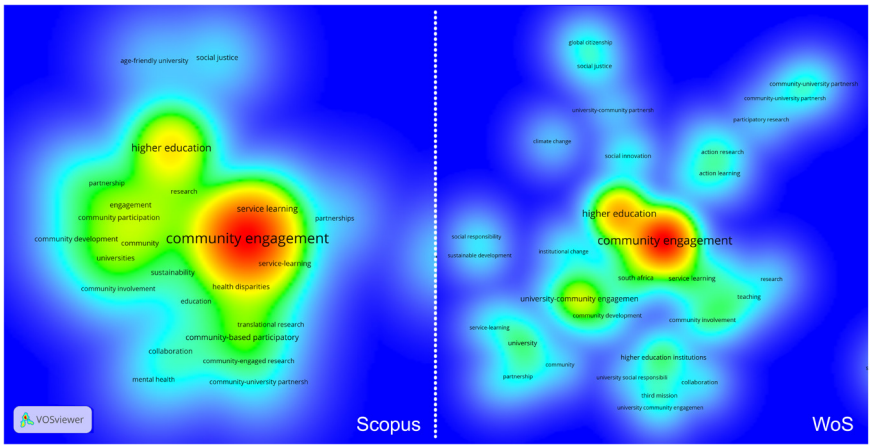
FIGURE VIII. Keywords most used by authors



Source: Compiled by author.

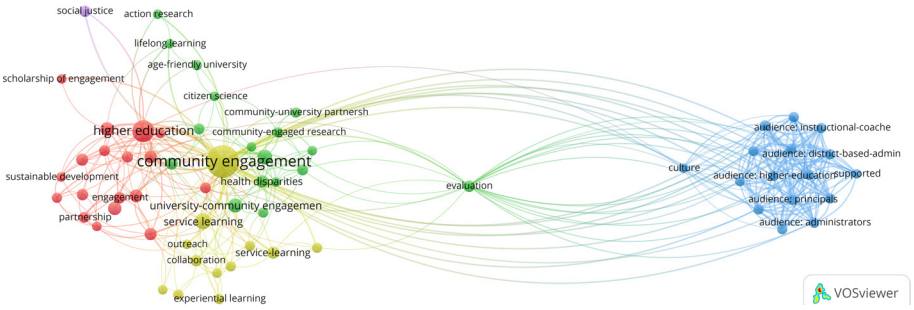
Received: 28-08-2022 Accepted: 11-05-2023

FIGURE IX. Keyword density visualization map according to bibliographic database



Source: Compiled by author.

FIGURE X. Keyword co-occurrence network

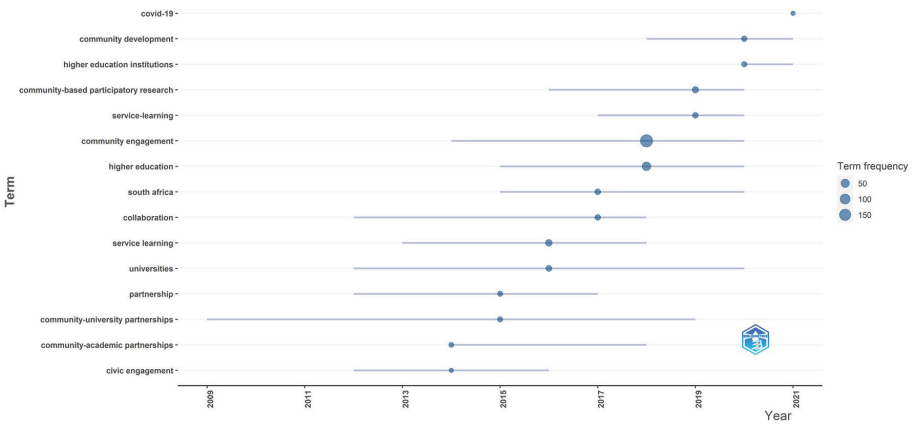


Source: Compiled by author.

Additionally, an examination of the dynamic evolution of keywords showed that before 2015, critical points on the subject were focused to basic themes about the link and partnership between academia and the community. After 2017, interest shifted towards more specific topics such as community development, community-based participatory research, service-learning, and more recently, the Covid-19 pandemic (Figure XI).

Received: 28-08-2022 Accepted: 11-05-2023

FIGURE XI. Keyword trend



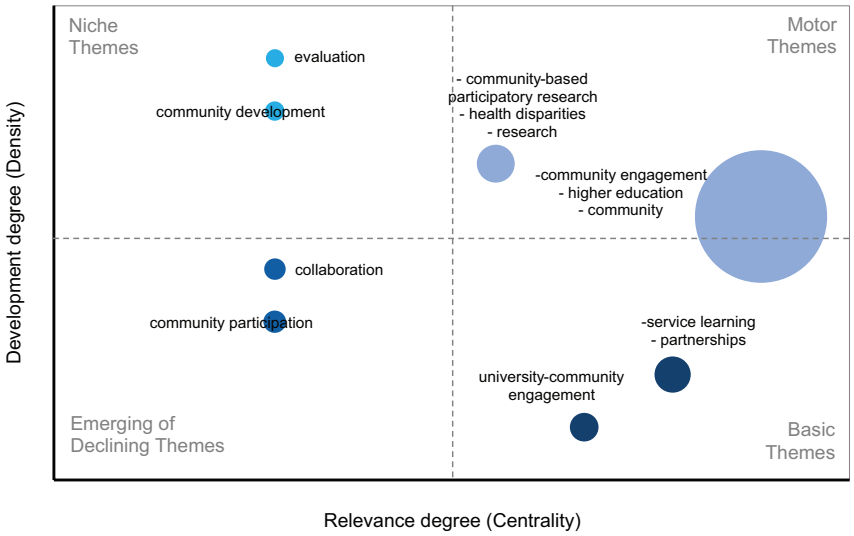
Source: Compiled by author.

Thematic analysis based on study categories

The algorithm of Callon et al. (1991) allowed to identify a total of eight clusters distributed between motor, basic, niche, and emerging themes (Figure XII).

The motor clusters were community engagement (189) and community-based participatory research (28). The first one grouped several case studies referring to the analysis and comparison of CE practices and policies in the world, applied in the context of higher education (Correa et al., 2012; Thomson et al., 2011; Winter et al., 2006). Others articles in this cluster address related topics such as University Social Responsibility (Chile & Black, 2015; Jones et al., 2021; Symaco & Tee, 2019) and Sustainable Development (Castro et al., 2021; Fissi et al., 2021; Mbah, 2018). The second cluster (which adds the concepts health disparities and research) grouped various publications associated with community-based participatory research (Johnston et al., 2019), its development complexities in particular areas such as health (Bodison et al., 2015; Weingart & Meyer, 2021), as well as reflections on the importance of institutionalizing CE as a university function (Bhagwan, 2020).

FIGURE XII. Thematic map of study categories



Source: Compiled by author.

The basic clusters were university-community engagement (12) and service-learning (20). The first one explored, in general, the theoretical-conceptual framework on CE (Nkoana & Dichaba, 2017; Pink et al., 2016). The works of Ehlenz (2018) and Bruning et al. (2006) which reviewed the role and contribution of universities for the communities in which they are inserted, stand out. The second cluster investigated, mostly, the service-learning methodology. Works that propose service-learning projects based on criteria of horizontality and bidirectionality (Gerstenblatt & Gilbert, 2014), guides for the design of community education initiatives led by students (Bunting et al., 2019), and pedagogical discussions on the application of service-learning actions (Smith et al., 2017) stand out.

The niche clusters were evaluation (6) and community development (8). The first cluster included research that presents the results and impacts of various educational initiatives linked to the CE (Mann & Bowen, 2021; Powers & Leili, 2017) and the study of key indicators for their measurement (Wanjiru & Xiaoguang, 2021). And, in the second cluster, publications that analyzed the importance of establishing the link

Received: 28-08-2022 Accepted: 11-05-2023

with the community engagement in stages of solutions design (Sheila et al., 2021) and the community impact in CE activities (Coetzee & Nell, 2018; Thomas et al., 2021) stand out.

The emerging or declining clusters were collaboration (10) and community participation (11). The first one, overall, explored the development of CE models to foster university-community collaboration (Bender, 2008; Bronstein et al., 2012; Swanberg et al., 2018). While the second covered a series of studies without major thematic integration, framed, however, in the field of mental health (Salzer et al., 2014, 2015; Snethen et al., 2021).

Discussion and Conclusions

The study of CE in universities is an increasingly important topic that is fundamentally related to sustainable development and the social responsibility of universities; topics that define the landscape of community participation transversely.

The literature review denotes a strong geographic concentration and publication sources, whose scientific production is distributed among a large number of authors, and is framed in the areas of education and health. Key concepts associated with community engagement and development, health care and empirical research in the clinical, medical and educational fields, were highlighted.

Over the last fifteen years, the study of CE has shown a consistent dynamic evolution, progressing from incipient research focused on exploring the university-community link, to more specific topics that highlight particular methodologies and practices for ensuring effective engagement.

The thematic analysis of the bibliography reveals that the catalysts of this field of study are the creation and development of sustainable partnerships with community stakeholders, and the role of universities in providing relevant community services and participatory research. This means that, as a general framework for their articulation with the community, universities must establish bidirectional and formalized practices aimed at increasing social capital for sustainable development through a participatory and timely model, as has been highlighted by other researchers (Bender, 2008; Chen, 2021; Zapp & Lerch, 2020).

In this way, it is suggests that CE must be based on a strategic and operational model of holistic approach, which integrates the university disciplines and functions to give it a sense and orientation towards the mission and educational project (Bender, 2008; Kruss, 2012). Likewise, sustainable development and university social responsibility are factors closely connected with CE and represent, together with the development of participatory teaching and learning methodologies, basic aspects for student engagement with the community.

In this context, it is not surprising that the mobilizing elements highlighted in the literature refer to the study of action research and community-based research processes, as well as the tools that allow for evaluating their results and impacts. It highlights, for example, the interest in the service learning (SL) methodology, an educational proposal in which participants learn while working on improving real needs of the environment.

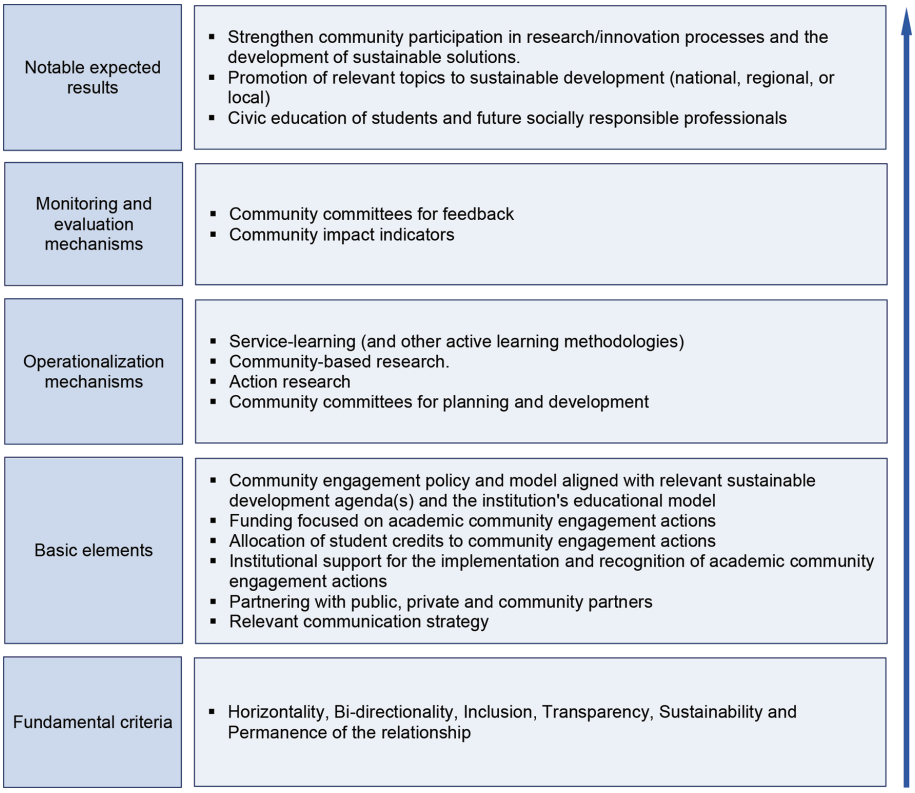
In conclusion, the results show that research within the framework of university CE has tended to raise the importance of this function as a strategic area of institutional management and a key aspect of academic management. Although it is still a flexible concept, there is a clear distinction between extension-volunteering actions and community engagement, because the latter must interact with other higher education functions. It is interesting to mention that although there is plenty of studies that scrutinize the university-community link, this is the first bibliometric review focused specifically on the notion of community engagement, which also includes a categorization based on the degree of interaction and strength of thematic association, facilitating new lines of research in the area. Thus, this research provides a general overview of the scientific production in CE in the context of university education.

For public policy makers, as well as academic managers, a general proposal for mechanisms for operationalization, monitoring and evaluation of the CE, synthesized in Figure XIII, is developed. It is important to note that the proposed mechanisms are built under a concrete policy framework that emphasizes the key roles of strategic planning, institutional leadership and effective control systems, and that ensure that the disclosure of information is open and transparent.

Finally, the study presents some limitations and opportunities for future research. From the perspective of bibliometric analysis, the review horizon can be expanded, as well as the databases consulted and the

Received: 28-08-2022 Accepted: 11-05-2023

FIGURE XIII. Proposal of mechanisms for operationalization, monitoring, and evaluation of CE



Source: Compiled by author.

analysis methods applied for investigating the structure and dynamics of research. Additionally, this exploration can be complemented and deepened with the investigation of the themes of the identified articles through, for example, the use of qualitative content analysis techniques. And from the results perspective, the findings show that although CE is a topic of general interest, its study focused on the field of health and medical sciences. Thus, the development of future research that considers other disciplinary areas such as social sciences, engineering and technology is encouraged. Also, it is relevant to strengthen the analysis

on the link with community partners (programs, practices, and methodologies), whose treatment would allow progress towards its consolidation as a driving issue. Similarly, it is considered important to propose research focused on evaluating the impact of university CE actions, currently classified as a niche topic. Furthermore, given the concentration of the spatial distribution of scientific production, it is necessary to generate initiatives in other geographical contexts (such as Latin America) that would significantly expand the panorama of research on CE.

All these opportunities will offer new theoretical and practical implications that, together with the current demands and challenges concerning the quality of higher education, will lead to new academic discussions that must be addressed in the future, safeguarding the scientific rigor revealed in this bibliometric examination on CE in the university context.

Bibliographical references

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Bender, G. (2008). Exploring conceptual models for community engagement at higher education institutions in South Africa. *Perspectives in Education, 26*(1), 81–95.

Bhagwan, R. (2020). Towards the institutionalisation of community engagement in higher education in South Africa. *Perspectives in Education, 38*(2), 36–55. https://doi.org/10.18820/2519593X/PIE.V38.I2.03

Bhatnagar, S., Agrawal, S., Sharma, D., & Singh, M. (2020). Perception on the Community Engagement of Students of Higher Education Institutions. *SSRN Electronic Journal*, 1–16. https://doi.org/10.2139/ssrn.3664052

Bidandi, F., Ambe, A. N., & Mukong, C. H. (2021). Insights and Current Debates on Community Engagement in Higher Education Institutions: Perspectives on the University of the Western Cape. *SAGE Open, 11*(2), 1–11. https://doi.org/10.1177/21582440211011467

Bodison, S. C., Sankaré, I., Anaya, H., Booker-Vaughns, J., Miller, A., Williams, P., & Norris, K. (2015). Engaging the Community in the Dissemination, Implementation, and Improvement of Health-Related Research. *Clinical and Translational Science, 8*(6), 814–819. https://doi.org/10.1111/cts.12342

Bradford, S. (1985). Sources of information on specific subjects 1934. *Journal of Information Science, 10*(4), 85–86. https://doi.org/10.1177/016555158501000407

Breznitz, S. M., & Feldman, M. P. (2012). The engaged university. *Journal of Technology Transfer, 37*(2), 139–157. https://doi.org/10.1007/s10961-010-9183-6

Bronstein, L. R., Anderson, E., Terwilliger, S. H., & Sager, K. (2012). Evaluating a model of school-based health and social services: An interdisciplinary community-university collaboration. *Children and Schools, 34*(3), 155–165. https://doi.org/10.1093/cs/cds004

Brown, K., Shephard, K., Warren, D., Hesson, G., & Fleming, J. (2016). Using phenomenography to build an understanding of how university people conceptualise their community-engaged activities. *Higher Education Research & Development, 35*(4), 643–657. https://doi.org/10.1080/07294360.2015.1137880

Bruning, S. D., McGrew, S., & Cooper, M. (2006). Town-gown relationships: Exploring university-community engagement from the perspective of community members. *Public Relations Review, 32*(2), 125–130. https://doi.org/10.1016/j.pubrev.2006.02.005

Bunting, S. R., Saqueton, R., & Batteson, T. J. (2019). A Guide for Designing Student-Led, Interprofessional Community Education Initiatives About HIV Risk and Pre-Exposure Prophylaxis. *MedEdPORTAL, 15*, 10818. https://doi.org/10.15766/mep_2374-8265.10818

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics, 22*(1), 155–205. https://doi.org/10.1007/BF02019280

Castro, A., Esmond, S., Perez, S., Scott, C., Jackson, T., & Passmore, S. R. (2021). Building Sustainable Organizational and Community Capacity for Research Partnerships: A Decade of Experience. *Progress in Community Health Partnerships: Research, Education, and Action, 15*(4), 553–560. https://doi.org/10.1353/cpr.2021.0045

Chen, A. S. (2021). Global Citizen Formation: Global Citizenship Education in Higher Education. In *Governance and Citizenship in Asia*. Springer. https://doi.org/10.1007/978-981-16-1959-5

Chile, L. M., & Black, X. M. (2015). University–community engagement: Case study of university social responsibility. *Education, Citizenship and Social Justice, 10*(3), 234–253. https://doi.org/10.1177/1746197915607278

Received: 28-08-2022 Accepted: 11-05-2023

Chung, B., Brown, A., Moreno, G., Cuen, P., Uy, V., Vangala, S., Bell, D., Washington, A., Norris, K., & Mangione, C. (2016). Implementing Community Engagement as a Mission at the David Geffen School of Medicine at the University of California, Los Angeles. *Journal of Health Care for the Poor and Underserved, 27*(1), 8–21. https://doi.org/10.1353/hpu.2016.0009

Chung, B., Norris, K., Mangione, C., del Pino, H. E., Jones, L., Castro, D., Wang, C., Bell, D., Vangala, S., Kahn, K., Brown, A. (2015). Faculty Participation in and Needs around Community Engagement within a Large Multiinstitutional Clinical and Translational Science Awardee. *Clinical and Translational Science, 8*(5), 506–512. https://doi.org/10.1111/cts.12314

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science Mapping Software Tools: Review, Analysis, and Cooperative Study Among Tools. *Journal of the American Society for Information Science and Technology, 62*(7), 1382–1402. https://doi.org/10.1002/asi.21525

Coetzee, H., & Nell, W. (2018). Measuring impact and contributions of South African universities in communities: The case of the North-West University. *Development Southern Africa, 35*(6), 1–17. https://doi.org/10.1080/0376835X.2018.1475218

Compagnucci, L., & Spigarelli, F. (2020). The Third Mission of the university: A systematic literature review on potentials and constraints. *Technological Forecasting and Social Change, 161*(March), 1–30. https://doi.org/10.1016/j.techfore.2020.120284

Correa, M. A., Butcher, J., & Howard, P. (2012). An international comparison of community engagement in higher education. *International Journal of Educational Development, 32*(1), 187–192. https://doi.org/10.1016/j.ijedudev.2011.04.008

Dave, G., Frerichs, L., Jones, J., Kim, M., Schaal, J., Vassar, S., Varma, D., Striley, C., Ruktanonchai, C., Black, A., Hankins, J., Lovelady, N., Cene, C., Green, M., Young, T., Tiwari, S., Cheney, A., Cottler, L., Sullivan, G., … Corbie-Smith, G. (2018). Conceptualizing trust in community-academic research partnerships using concept mapping approach: A multi-CTSA study. *Evaluation and Program Planning, 66*(2), 70–78. https://doi.org/10.1016/j.evalprogplan.2017.10.007

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*(May), 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Dougnac, P. (2016). Una revisión del concepto anglosajón public engagement y su equivalencia funcional a los de extensión y vinculación con el medio. *Pensamiento Educativo: Revista de Investigación Educacional Latinoamericana, 53*(2), 1–19. https://doi.org/10.7764/PEL.53.2.2016.11

Dry, S., Garrett, S., Koenig, B., Brown, A., Burgess, M., Hult, J., Longstaff, H., Wilcox, E., Contreras, S., Martinez, A., Boyd, E., & Dohan, D. (2017). Community recommendations on biobank governance: Results from a deliberative community engagement in California. *PLoS ONE, 12*(2), e0172582. https://doi.org/10.1371/journal.pone.0172582

Ebersohn, L., Loots, T., Eloff, I., & Ferreira, R. (2015). Taking note of obstacles research partners negotiate in long-term higher education community engagement partnerships. *Teaching and Teacher Education, 45*(1), 59–72. https://doi.org/10.1016/j.tate.2014.09.003

Ehlenz, M. M. (2018). Defining University Anchor Institution Strategies: Comparing Theory to Practice. *Planning Theory and Practice, 19*(1), 74–92. https://doi.org/10.1080/14649357.2017.1406980

Fissi, S., Romolini, A., Gori, E., & Contri, M. (2021). The path toward a sustainable green university: The case of the University of Florence. *Journal of Cleaner Production, 279*, 1–9. https://doi.org/10.1016/j.jclepro.2020.123655

Forliano, C., De Bernardi, P., & Yahiaoui, D. (2021). Entrepreneurial universities: A bibliometric analysis within the business and management domains. *Technological Forecasting and Social Change, 165*(1–16). https://doi.org/10.1016/j.techfore.2020.120522

Gerstenblatt, P., & Gilbert, D. J. (2014). Framing Service Learning in Social Work: An Interdisciplinary Elective Course Embedded within a University–Community Partnership. *Social Work Education, 33*(8), 1037–1053. https://doi.org/10.1080/02615479.2014.935731

Gilmore, B., Ndejjo, R., Tchetchia, A., De Claro, V., Mago, E., Diallo, A. A., Lopes, C., & Bhattacharyya, S. (2020). Community engagement for COVID-19 prevention and control: A rapid evidence synthesis. *BMJ Global Health, 5*(10), 1–11. https://doi.org/10.1136/bmjgh-2020-003188

Govender, I., & Giampiccoli, A. (2020). Towards university involvement in community-based tourism monitoring & evaluation education. *e-Review of Tourism Research, 18*(2), 252–268.

Greenhill, J. A., Walker, J., & Playford, D. (2015). Outcomes of Australian Rural Clinical Schools: A Decade of Success Building the Rural Medical

Workforce Through the Education and Training Continuum. *Rural and Remote Health, 15*(3), 1–14. https://doi.org/10.22605/RRH2991

Groark, C. J., & McCall, R. B. (2018). Lessons learned from 30 years of a university-community engagement center. *Journal of Higher Education Outreach and Engagement, 22*(2), 7–30.

Johnston, A., Wells, P., & Woodhouse, D. (2021). Examining the roles of universities in place-based industrial strategy: which characteristics drive knowledge creation in priority technologies? *Regional Studies*, 1–12. https://doi.org/10.1080/00343404.2021.1956683

Johnston, J., Lopez, M., Gribble, M. O., Gutschow, W., Austin, C., & Arora, M. (2019). A Collaborative Approach to Assess Legacy Pollution in Communities Near a Lead–Acid Battery Smelter: The "Truth Fairy" Project. *Health Education and Behavior, 46*(1S), 71S-80S. https://doi.org/10.1177/1090198119859406

Johnston, K., & Lane, A. (2018). Building relational capital: The contribution of episodic and relational community engagement. *Public Relations Review, 44*(5), 633–644. https://doi.org/10.1016/j.pubrev.2018.10.006

Jones, E., Leask, B., Brandenburg, U., & de Wit, H. (2021). Global Social Responsibility and the Internationalisation of Higher Education for Society. *Journal of Studies in International Education, 25*(4), 330–347. https://doi.org/10.1177/10283153211031679

Kruss, G. (2012). Reconceptualising engagement: A conceptual framework for analysing university interaction with external social partners. *South African Review of Sociology, 43*(2), 5–26. https://doi.org/10.1080/21528586.2012.694240

Kruss, G., Haupt, G., & Visser, M. (2016). 'Luring the academic soul': promoting academic engagement in South African universities. *Higher Education Research & Development, 35*(4), 755–771. https://doi.org/10.1080/07294360.2015.1137874

Lavery, J. V., Tinadana, P. O., Scott, T. W., Harrington, L. C., Ramsey, J. M., Ytuarte-Nuñez, C., & James, A. A. (2010). Towards a framework for community engagement in global health research. *Trends in Parasitology, 26*(6), 279–283. https://doi.org/10.1016/j.pt.2010.02.009

Levin, M. B., Bowie, J. V., Ragsdale, S. K., Gawad, A. L., Cooper, L. A., & Sharfstein, J. M. (2021). Enhancing Community Engagement by Schools and Programs of Public Health in the United States. *Annual Review of Public Health, 42*, 405–421. https://doi.org/10.1146/annurev-publhealth-090419-102324

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences, 16*(12), 317–323.

Machimana, E., Ebersöhn, L., & Sefotho, M. (2018). What makes or breaks higher education community engagement in the South African rural school context: A multiple-partner perspective. *Education Citizenship and Social Justice, 13*(2), 177–196. https://doi.org/10.1177/1746197917731353

Machimana, E., Ebersöhn, L., & Sefotho, M. (2020). Higher education community engagement as a pathway to developing global citizenship practices in young people South African perspective. *Citizenship Teaching and Learning, 15*(3), 371–387.

Machimana, E. G., Sefotho, M. M., Ebersöhn, L., & Shultz, L. (2021). Higher education uses community engagement-partnership as a research space to build knowledge. *Educational Research for Policy and Practice, 20*(1), 45–62. https://doi.org/10.1007/s10671-020-09266-6

Mann, J., & Bowen, G. A. (2021). Emphasizing the Impact of Service-Learning and Community Engagement Partnerships Amid Challenges Facing Higher Education. *International Journal for Research on Service-Learning and Community Engagement, 9*(1), 1–5. https://doi.org/10.37333/001C.31379

Mbah, M. (2018). Can local knowledge make the difference? Rethinking universities' community engagement and prospect for sustainable community development. *The Journal of Environmental Education*, 1–12. https://doi.org/10.1080/00958964.2018.1462136

Mbah, M., Johnson, A. T., & Chipindi, F. M. (2021). Institutionalizing the intangible through research and engagement: Indigenous knowledge and higher education for sustainable development in Zambia. *International Journal of Educational Development, 82*, 1–9. https://doi.org/10.1016/j.ijedudev.2021.102355

McNaughton, D., & Duong, T. T. H. (2014). Designing a Community Engagement Framework for a New Dengue Control Method: A Case Study from Central Vietnam. *PLoS Neglected Tropical Diseases, 8*(5), 1–12. https://doi.org/10.1371/journal.pntd.0002794

Mitchell, N., Triska, M., Liberatore, A., Ashcroft, L., Weatherill, R., & Longnecker, N. (2017). Benefits and challenges of incorporating citizen science into university education. *PLoS ONE, 12*(11), 1–15. https://doi.org/10.1371/journal.pone.0186285

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7). https://doi.org/10.1371/journal.pmed.1000097

Mtawa, N. N., Fongwa, S. N., & Wangenge-Ouma, G. (2016). The scholarship of university-community engagement: Interrogating Boyer's model. *International Journal of Educational Development, 49*, 126–133. https://doi.org/10.1016/j.ijedudev.2016.01.007

Muñoz-Leiva, F., Viedma-del-Jesús, M. I., Sánchez-Fernández, J., & López-Herrera, A. G. (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality and Quantity, 46*, 1077–1095. https://doi.org/10.1007/s11135-011-9565-3

Mutero, I., & Govender, I. (2019). Moving from Transactional Partnerships to Collaborative University Community Engagement: A Case Study Evaluating Creative Placemaking in Kwazulu-Natal Province. *South African Review of Sociology, 50*(1), 3–17. https://doi.org/10.1080/21528586.2019.1632578

Mutero, I., & Govender, I. (2020). Advancing the Exploration of Engaged Creative-Placemaking Amongst Universities and Communities for Social Cohesion in South Africa. *Journal of Asian and African Studies*, *55*(3), 429–445. https://doi.org/10.1177/0021909619882718

Mutero, I., & Govender, I. (2021). Monitoring and Evaluation of Transformative Creative Placemaking on University Campuses. *Journal of Asian and African Studies, 56*(7), 1577–1589. https://doi.org/10.1177/0021909620975804

Nkoana, E. M., & Dichaba, M. M. (2017). Development and application of conceptual and analytic frameworks for community engagement at a South African higher education institution. *South African Journal of Higher Education, 31*(6), 177–196. https://doi.org/10.20853/31-6-1574

Osborne, J., Paget, J., Giles-Vernick, T., Kutalek, R., Napier, D., Baliatsas, C., & Dückers, M. (2021). Community engagement and vulnerability in infectious diseases: A systematic review and qualitative analysis of the literature. *Social Science and Medicine, 284*(July), 1–12. https://doi.org/10.1016/j.socscimed.2021.114246

Petersen, I., & Kruss, G. (2021). Universities as change agents in resource-poor local settings: An empirically grounded typology of engagement

models. *Technological Forecasting and Social Change, 167*(2), 120693. https://doi.org/10.1016/j.techfore.2021.120693

Pink, M. A., Taouk, Y., Guinea, S., Bunch, K., Flowers, K., & Nightingale, K. (2016). Developing a conceptual framework for student learning during international community engagement. *Journal of University Teaching and Learning Practice, 13*(5).

Pommier, E., Neff, K. D., & Tóth-Király, I. (2020). The Development and Validation of the Compassion Scale. *Assessment, 27*(1), 21–39. https://doi.org/10.1177/1073191119874108

Powers, R. A., & Leili, J. (2017). Bar Training for Active Bystanders: Evaluation of a Community-Based Bystander Intervention Program. *Violence Against Women*, 1–21. https://doi.org/10.1177/1077801217741219

Richardson, E. Z. L., Bandewar, S. V. S., Boulanger, R. F., Mehta, R., Lin, T., Vincent, R., Molyneux, S., Goldstone, A., & Lavery, J. V. (2021). Addressing diversity and complexity in the community engagement literature: The rationale for a realist review. *Wellcome Open Research, 5*(1), 1–16. https://doi.org/10.12688/wellcomeopenres.15525.2

Robertson, S. L. (2021). Global competences and 21st century higher education – And why they matter. *International Journal of Chinese Education, 10*(1), 1–9. https://doi.org/10.1177/22125868211010345

Ryan, B., Johnston, K. A., Taylor, M., & McAndrew, R. (2020). Community engagement for disaster preparedness: A systematic literature review. *International Journal of Disaster Risk Reduction, 49*(March), 1–13. https://doi.org/10.1016/j.ijdrr.2020.101655

Salzer, M. S., Brusilovskiy, E., Prvu-Bettger, J., & Kottsieper, P. (2014). Measuring community participation of adults with psychiatric disabilities: Reliability of two modes of data collection. *Rehabilitation Psychology, 59*(2), 211–219. https://doi.org/10.1037/a0036002

Salzer, M. S., Kottsieper, P., & Brusilovskiy, E. (2015). Intermethod reliability and factors affecting recall with the Temple University Community Participation measure. *Journal of Mental Health, 24*(4), 189–195. https://doi.org/10.3109/09638237.2015.1036976

Sheila, N. A., Zhu, C., Kintu, M. J., & Kataike, J. (2021). Assessing higher education institutional stakeholders' perceptions and needs for community engagement: An empirical evidence from Uganda. *Heliyon, 7*(4), 1–9. https://doi.org/10.1016/j.heliyon.2021.e06612

Shephard, K., Brown, K., & Guiney, T. (2017). Researching the Professional-Development Needs of Community-Engaged Scholars in a New Zealand University. *Sustainability, 9*(7), 1249. https://doi.org/10.3390/su9071249

Shephard, K., Brown, K., Guiney, T., & Deaker, L. (2018). Valuing and evaluating community-engaged scholarship. *Tertiary Education and Management, 24*(1), 83–94. https://doi.org/10.1080/13583883.2017.1395904

Skute, I., Zalewska-Kurek, K., Hatak, I., & de Weerd-Nederhof, P. (2019). Mapping the field: a bibliometric analysis of the literature on university–industry collaborations. *Journal of Technology Transfer, 44*(3), 916–947. https://doi.org/10.1007/s10961-017-9637-1

Smith, B. D., Marshall, I., Anderson, B. E., & Daniels, K. K. (2017). A partnership forged: BSW students and service learning at a historically Black college and university (HBCU) serving urban communities. *Journal of Human Behavior in the Social Environment, 27*(5), 438–449. https://doi.org/10.1080/10911359.2017.1295005

Snethen, G., Brusilovskiy, E., McCormick, B. P., Hiremath, S. V., & Salzer, M. S. (2021). The relationship between community participation and physical activity among individuals with serious mental illnesses. *Mental Health and Physical Activity, 20*(2021), 1–8. https://doi.org/10.1016/j.mhpa.2021.100381

Swanberg, S. M., Mi, M., Engwall, K., & Bulgarelli, N. (2018). Community Engagement at an Emerging Academic Medical Library: A Three-Pronged Outreach Model. *Medical Reference Services Quarterly, 37*(3), 234–248. https://doi.org/10.1080/02763869.2018.1477707

Symaco, L. P., & Tee, M. Y. (2019). Social responsibility and engagement in higher education: Case of the ASEAN. *International Journal of Educational Development, 66*, 184–192. https://doi.org/10.1016/j.ijedudev.2018.10.001

Thomas, E., Snethen, G., & Salzer, M. (2017). A developmental study of community participation of individuals with serious mental illnesses: Implications for policy and practice. *American Journal of Orthopsychiatry, 87*(5), 597–605. https://doi.org/10.1037/ort0000269

Thomas, T., Joseph, M. K., Nazar, L., Ravi, S., & Vijesh, P. V. (2021). Engaging in Community Development Practices, Public Libraries as Flagship Bearers: A case study of Mulanthuruthy Public Library in Ernakulum District, Kerala, India. *Library Philosophy and Practice*, 1–11.

Thomson, A. M., Smith-Tolken, A. R., Naidoo, A. V., & Bringle, R. G. (2011). Service Learning and Community Engagement: A Comparison of Three National Contexts. *Voluntas, 22*(2), 214–237. https://doi.org/10.1007/s11266-010-9133-9

Unger, M., & Polt, W. (2017). The knowledge triangle between research, education and innovation -A conceptual discussion. *Foresight and STI Governance, 11*(2), 10–26. https://doi.org/10.17323/2500-2597.2017.2.10.26

UNICEF. (2020). *Minimum Quality Standards and Indicators for Community Engagement*. UNICEF.

Van Eeden, E., Eloff, I., & Dippenaar, H. (2021). On Responses of Higher Education and Training With(in) Society Through Research, Teaching and Community Engagement. *Educational Research for Social Change, 1*0(1), 1–15.

Wanjiru, I. R., & Xiaoguang, L. (2021). Evaluating University–Community Engagement Through a Community-Based Lens: What Indicators Are Suitable? *Journal of Higher Education Outreach and Engagement, 25*(4), 133–152.

Weerts, D. (2014). State Funding and the Engaged University: Understanding Community Engagement and State Appropriations for Higher Education. *The Review of Higher Education, 38*(1), 133–169. https://doi.org/10.1353/rhe.2014.0038

Weerts, D. (2016). From Covenant to Contract: Changing Conceptions of Public Research Universities in American Society. *The Good Society, 25*(2-3), 182–208. https://doi.org/10.5325/goodsociety.25.2-3.0182

Weerts, D. J., & Sandmann, L. R. (2008). Building a two-way street: Challenges and opportunities for community engagement at research universities. *Review of Higher Education, 32*(1), 73–106. https://doi.org/10.1353/rhe.0.0027

Weerts, D. J., & Sandmann, L. R. (2010). Community engagement and boundary-spanning roles at research universities. *The Journal of Higher Education, 81*(6), 632–657. https://doi.org/10.1080/00221546.2010.11779075

Weingart, P., & Meyer, C. (2021). Citizen science in South Africa: Rhetoric and reality. *Public Understanding of Science, 30*(5), 1–16. https://doi.org/10.1177/0963662521996556

Winter, A., Wiseman, J., & Muirhead, B. (2006). University-community engagement in Australia: Practice, policy and public good. *Education, Citizenship and Social Justice, 1*(3), 211–230. https://doi.org/10.1177/1746197906064675

Yuan, M., Lin, H., Wu, H., Yu, M., Tu, J., & Lü, Y. (2021). Community engagement in public health: a bibliometric mapping of global research. *Archives of Public Health, 79*(1), 1–17. https://doi.org/10.1186/s13690-021-00525-3

Zapp, M., & Lerch, J. C. (2020). Imagining the World: Conceptions and Determinants of Internationalization in Higher Education Curricula Worldwide. *Sociology of Education, 93*(4), 372–392. https://doi.org/10.1177/0038040720929304

Zavale, N. C., & Langa, P. V. (2018). University-industry linkages' literature on Sub-Saharan Africa: systematic literature review and bibliometric account. *Scientometrics, 116*(1), 1–49. https://doi.org/10.1007/s11192-018-2760-4

Zhu, J., & Liu, W. (2020). A tale of two databases: the use of Web of Science and Scopus in academic papers. *Scientometrics, 123*(2), 321–335. https://doi.org/10.1007/s11192-020-03387-8

Zupic, I., & Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods, 18*(3), 429–472. https://doi.org/10.1177/1094428114562629

Contact address: Sebastián Araya-Pizarro. Universidad de La Serena, Facultad de Ciencias Sociales, Empresariales y Jurídicas, Departamento de Ciencias Económicas y Empresariales, Amunátegui s/n, La Serena, Chile. E-mail: saraya@userena.cl

Received: 28-08-2022 Accepted: 11-05-2023

Female Pedagogical Leadership in Latin American Schools

Liderazgo Pedagógico Femenino en Escuelas de América Latina

https://doi.org/10.4438/1988-592X-RE-2023-402-599

Amparo Jiménez Vivas
https://orcid.org/0000-0002-2739-6581
Universidad Pontificia de Salamanca

Patricia Lorena Parraguez Núñez
https://orcid.org/0000-0001-6554-3338
Universidad Pontificia de Salamanca

Abstract

In the complex framework that characterizes Latin American education, marked by strong inequality and unsatisfactory results further intensified by the COVID-19 pandemic, this article describes and analyzes, from a gender perspective, the pedagogical leadership of a group of Salesian Catholic schools in 21 countries of the continent, considered by its relevance as a major factor towards improving learning. For this purpose, a quantitative study was carried out by applying the questionnaire technique, with a probabilistic and stratified sample of 330 principals. Among the main results, we found that there was a predominance of female school leadership (66.7%) and more consolidated pedagogical approach practices in women who perform this role, such as supervision and monitoring of teaching practice, collaborative culture for professional development and decision making oriented to teaching and learning. There are other practices that require greater drive and training for all principals and their teams.

Keywords: gender issues, educational administration, Catholic schools, instructional leadership, statistical analysis.

Resumen

En el complejo entramado que caracteriza la educación latinoamericana, marcada por una fuerte desigualdad y resultados insatisfactorios aún más acentuados con la pandemia de la COVID-19, este artículo describe y analiza, desde una perspectiva de género, el liderazgo pedagógico de un conjunto de escuelas católicas salesianas de 21 países del continente, considerado por su relevancia como un factor de primer orden para la mejora de los aprendizajes. Para ello se realizó una investigación de carácter cuantitativo a través de la aplicación de la técnica del cuestionario, con una muestra probabilística y estratificada que abarcó a 330 directores/as. Entre los principales resultados encontramos el predominio del liderazgo escolar femenino (66.7%) y prácticas de enfoque pedagógico más consolidadas en las mujeres que ejercen este rol, tales como: supervisión y monitoreo de la práctica docente, cultura colaborativa para el desarrollo profesional y toma de decisiones orientada a la enseñanza y el aprendizaje. Se constatan otras prácticas que requieren mayor impulso y formación en el conjunto de los directores y sus equipos.

Palabras claves: mujer y desarrollo, gestión educacional, escuelas confesionales, liderazgo pedagógico, análisis bivariado.

Introduction

The purpose of this research is to have an approach to the characteristics of pedagogical leadership exercised by people in leadership positions in Salesian schools in America, considering the relevance achieved by this approach and its relevance for the Latin American context (Bolivar, 2010; Murillo, 2007; Pont et al., 2008). For Salesian schools, the challenge of overcoming inequalities and narrowing educational gaps throughout the continent leads them to initiate research processes regarding their extensive presence in 25 countries of the Americas and the factors that influence the quality of education. Among these, school leadership is one of the most important factors, especially when it contributes to the transformation of educational organizations into learning communities (Montecinos and Cortez, 2015).

In the same vein, this study has placed significant value on female school leaders, since their capacity to influence and practice leadership in schools in the Americas accounts for a much higher percentage than that of men (Weinstein et al., 2014). In turn, several international studies on women and educational leadership acknowledge the need to continue

Received: 07-12-2022 Accepted: 05-05-2023

including the gender variable in research on this topic, to shed light on the possible symptoms of discrimination and inequality faced by women in the workplace, in order to implement solutions to the problem (Cárdenas de Sanz, 2017; Castro et al., 2021; Rivera-Mata, 2013).

In the relentless quest to achieve good student learning, and after decades of promoting major external educational reforms, there is now consensus on the limited capacity they have had to significantly transform the educational reality (Fullan 2002; Loyo, 2019). On the contrary, the capacity to make their own decisions as a school and to establish pedagogical direction become powerful devices for schools to provide organizational responses in accordance with the demands of their respective contexts. Hence, the educational or pedagogical leadership of school management and management teams acquires the character of a critical factor of the first order (Bolívar, 2019).

The multifaceted educational scenario in Latin America has been characterized by a search for practices that make it possible to achieve quality learning, thus allowing students to overcome the margins of inequality so deeply rooted in the continent. The recognition that school leaders play a highly relevant role in raising the quality of teaching and learning has generated significant interest in offering adequate preparation for the exercise of school management leadership, based on policies aimed at strengthening school leaders (Weinstein et al., 2015). Several innovative programs that have been developed in Latin American countries have acquired new leadership approaches, such as pedagogical, transformational, and distributed leadership, while moving from a leadership focused on administration or a bureaucratic approach to schools to a leadership for change and continuous improvement (Murillo, 2006).

On the other hand, it is noted that female leadership is undergoing major paradigm shifts. Women have been positioning themselves as capable of making decisions in various areas of their personal and professional lives, as well as performing in multiple areas (Güezmes et al., 2022). However, in relation to the educational scenario, various studies confirm that, despite the progress and feminization of education, women's entry into management positions is lower compared to men (Carrillo, 2017; Cruz-González et al., 2020; Padilla, 2018).

In Latin America, however, the evidence showing a high presence of women in management positions is consistent (Murillo and Román, 2013; Weinstein et al., 2014), which some identify as a "continent of

empowered women" (Cárdenas de Sanz, 2017:13), thus corroborating that female leadership has a broad capacity to face challenges, especially in contexts of change, encouraging flexibility and adaptability in the face of continuous transformations (Navarro et al., 2018).

The relevance of pedagogical leadership for improving the quality of schools in Latin America supports the importance and the need to continue to deepen this topic from a gender perspective, which although the amount of research in the region has increased, it is still insufficient (García and Martínez, 2019).

Conceptual Framework

- Approaches to Pedagogical Leadership

Pedagogical leadership has been developed by two major traditions with diverse approaches and objectives. The first refers to a type of "instructional leadership" of North American origin (Bush and Glover, 2014; Coughlin and Baird, 2013; Farnsworth, 2015), which has assumed that the critical focus of attention on the part of leaders is the behavior of teachers and their relationship to activities that directly affect student learning; the other, in contrast, that of "learning-centered leadership" (Hallinger and Heck, 2010; Leithwood et al., 2006; Robinson, 2011; Robinson et al., 2009), integrates instructional leadership with transformational leadership, which focuses on incorporating a broad spectrum of leadership actions to sustain learning and its outcomes (Hallinger, 2010; Lewis and Murphy, 2008).

Currently, school leadership is gaining momentum as leadership for learning or learning-centered leadership (learning-centered leadership). In the words of Robinson (2017:45) "the essence of learner-centered leadership is the permanent focus on the consequences that the decisions and actions of leaders have on the students for whom they are responsible." That is, a critical connection is made between the exercise of leadership with student learning. This could be obvious, since the purpose of schools is precisely learning and school leadership should be based on it; however, we can question why something that should be obvious is novel in today's educational contexts. And one possible answer is that leadership has not been directly related to student learning outcomes since it is usually an individual and independent responsibility of teachers and their

Received: 07-12-2022 Accepted: 05-05-2023

work in the classroom. That is, school leadership has been radically disconnected from teaching and learning for too long (Robinson, 2011: 8).

Studies on effective schools have systematically shown that one of the main characteristics that make them effective concerns the school director as pedagogical leader. Subsequently, authors such as Barber and Mourshed (2007), Bolívar and Murillo (2017) and Leithwood et al. (2008), have found that the school principal, when exercising pedagogical leadership functions, is the second factor, after classroom teaching, which has the greatest impact on student learning within intra-school factors.

In a study developed by Murillo (2007) in eight Latin American countries, it was found that there is a statistically significant relationship between the time principals devote to tasks related to pedagogical leadership and better performance of students in that school, and between this dedication and the greater satisfaction of teachers with the directorate.

Therefore, school leadership focused on pedagogical (Bolívar, 2012; Leithwood, 2009) and continuous improvement is proving to be an important way to face dilemmas and challenges of increasing complexity, providing a new theoretical lens that makes it possible to reconceptualize and reconfigure the practice of leadership in schools (Murillo, 2006). In the scenario of the health crisis, leading with a focus on teaching and learning has constituted an important demand for the leadership of the system, having to maintain an institutional management attentive to multiple demands.

■ Female Leadership in Education

The Latin American region reveals a high percentage of women linked to education in the Americas, and that has boosted the participation rate in school leadership, above the average of women in management positions in other spheres of social life (Cárdenas de Sanz, 2017).

In recent years, research on female leadership has increased significantly with the aim of evidencing the barriers that women face when facing management positions in the school scenario (Cruz et al., 2020; Cuevas et al., 2014), among which stand out experiences of successful women who are bringing greater visibility to important leadership achievements in large jobs (Malcorra, 2018; Sandberg and Scovell, 2015).

Some reasons that explain the increase of female leadership and positions of power in academia relate to the international agenda of gender

equity and female empowerment. However, it is no less true that in these processes marked actions of discrimination in terms of managerial permanence coexist, due to the enduring stereotype of this position, linked more to male optics. This conditions women to reproduce that role (Navarro et al., 2018; Rivera-Mata, 2013).

Female leadership in educational institutions is linked to the involvement of all entities responsible for the organizational system, transferring the commitment obtained around the quality and culture of the organization, i.e., leadership of orientation, motivation, development of empathy, and interpersonal skills. It also includes optimal educational management and its facilitating projection in the community (Martínez and Martínez, 2012). The contribution generated from this type of leadership has shown that women are capable of exercising leadership at different levels, which is a way to move towards breaking hierarchical structures that hinder change in schools and in processes outside them, which is an important contribution to the challenges of the current reality (Cáceres et al, 2012).

According to Bolívar (2019), pedagogical leadership drives the creation of necessary conditions for learning, with a focus on curriculum, pedagogy, providing the educational team with ambitious learning goals for students and school autonomy. Pedagogical leadership is characterized by the importance it gives to teaching practice, focusing on teaching and evaluation, as well as on the professional development of teachers (Montecinos and Cortés, 2015), and in the search for building the best conditions for learning processes; promoting a collaborative culture, minimizing individualistic glimpses in teaching practices (Llorent-Bedmar et al., 2017) and ensuring that decisions regarding management are always driven by teaching and learning (Bendikson et al., 2012).

A study by Carrasco and Barraza (2021) shows that the pedagogical leadership approach, with its peculiarities, is linked to female leadership. It also states that female leaders characteristically work better with others, tending to develop more horizontal and collaborative work environments (Kaiser and Wallace, 2016). However, this approach would not be the only one, as female leadership has several particularities of other types of leadership, such as transformational, distributed, or leadership for social justice. In other words, we are gradually moving from the old concept of "think manager: think man" (Schein, 1973) to revaluating

the contribution that female leadership can make to organizations. As various theories and studies on leadership have been developed, it has been shown that female leadership can combine effectiveness, concern for people, influence, inspirational motivation, adaptation to the diverse needs of the context, and cooperation more easily than men (Omar and Davidson, 2001; Wajcman, 1996).

If we currently understand that leadership is impossible without interpersonal influence and group effectiveness, the main focus of leadership today is on achieving the effective relationship between members, evolving from an individual to a collective construction (Rivera-Mata, 2013). In other words, in the social environment of the 21st century, there is no meaningful leadership without the relationships between people, or without considering the various contexts in which leadership unfolds, which are increasingly complex, changing, and interdependent.

■ Salesian School in the Americas and female school leadership

Taking Salesian schools in the Southern Cone as a reference, a study conducted in 2020 (Jiménez et al., 2023)) reveals that out of a total of 300 principals surveyed, 67.3% were women and 32.7% were men. In other words, in the 219 schools that make up this region (Argentina, Chile, Paraguay, and Uruguay), a high percentage is made up of women in management positions. This choice does not seem to be accidental. For the Congregation of the Daughters of Mary Help of Christians, founded by St. John Bosco and St. Mary Mazzarello (1872) in northern Italy and present in 97 countries, the formation and promotion of women is a priority (FMA, 1996; 1997). The Salesian Congregation, for its part, in recent decades has given significant space to the role of women in the educational community (Salesian Congregation, 1996), providing them with opportunities for their professional and personal development. There is also particular commitment of female teachers prepare themselves professionally to take on new challenges, which makes promotion to leadership positions frequent and natural in Salesian schools in America, along with other gender characteristics that are expressed with greater visibility by female leadership, such as the ability to work with others, to promote participatory decisions, sensitivity to their own emotional world and that of others, and passion for education from the ethics of care, that is, a people-centered leadership (Arroyo, 2020; Brescoll, 2016).

■ Pedagogical leadership and pandemic

With the changes in the scenario brought about by Covid-19, the school and its leaders have had to adapt to a new work model, to new challenges and transformations that reveal the need to reflect on the future. Thus, leaders must be able to engage their teams and the entire school community, to take on the challenge of reconfiguring the school, redesigning the conditions to ensure it is implemented.

In relation to women's leadership in times of pandemic, the UN states that "the leadership style of women leaders in the response to Covid-19 has been described as more collective than individual, more collaboration than competition and more directive than imperative" (Zedník, 2020), important characteristics in a context where the conception of leadership involves resilience, mobilizing human and material resources to achieve shared goals, aggregating capabilities to respond to unexpected events, and staying the pedagogical course in a remote, hybrid, or face-to-face learning ecosystem (Vaillant, 2022).

Although no specific analysis has been found on female pedagogical leadership in times of pandemic, it may be inferred that reacting to the emergency could have been faster, more decisive and effective, due to the fact that female characteristics are more suitable for times of crisis, as demonstrated by the theory of "think crisis - think female" (Gartzia et al., 2012); or as referred by the study of Zenger and Folkman (2021), where women tended to perform better during the crisis and were positively rated with higher statistical significance than men in the first wave of the pandemic.

Objectives

The objective of this article is to describe and analyze from a gender perspective, the pedagogical leadership that is developed in Salesian schools in the Americas, in order to characterize the reality of the continent and to promote further formative processes of school leaders. Based on the data obtained, it particularly intends to:

a) Describe the results of the pedagogical leadership variables that were addressed.

b) Determine whether there are significant differences in the exercise of pedagogical leadership based on the gender variable.

Received: 07-12-2022 Accepted: 05-05-2023

c) To recognize the practices most consolidated by male and female director in the continent and those that need to be strengthened.

Method

The article is based on the preliminary findings of a doctoral research project on school leadership in Salesian schools in Latin America. This project seeks to improve the current understanding of school leadership practices in this region. The quantitative method has been chosen, through a non-experimental and cross-sectional design.

Sample

The study was carried out between April and September 2021, with the participation of 330 principals of Salesian schools, located in urban and rural areas. The population corresponds to a set of about 1000 educational centers in 21 Latin American countries, including the United States only as a general reference.

A random, probabilistic sample was determined, which made it possible to extrapolate and thus generalize the results observed to the accessible population; and from there, to the general population. The type of sampling was also stratified, which was constructed based on the countries that constitute the universe and the number of Salesian schools that exist in each of them, thus forming strata, which were randomly selected through the SPSS statistical program, in proportion to the universe present in each country. Since the population was finite in size, the random sample was calculated considering a correction factor, an estimated error of ±5%, and a 95% confidence level.

Instrument

The instrument selected for data collection was the Principal Instructional Leadership Rating Scale (PIMRS) by Philip Hallinger (2015), which has been widely used in several studies and doctoral theses. An adaptation

Received: 07-12-2022 Accepted: 05-05-2023

was made for Latin American Spanish and Portuguese (Brazil), with expert validation. The questionnaire was applied online through the LimeSurvey platform, with ten variables associated with pedagogical leadership (Table I), each with 5 Likert-type questions and five levels of responses: (1) Almost never; (2) Rarely; (3) Sometimes; (4) Frequently; (5) Almost always.

TABLE I. Variables Associated with Pedagogical Leadership

DEFINE THE SCHOOL'S MISSION	MANAGE THE INSTRUCTIONAL PROGRAM	DEVELOP THE SCHOOL'S LEARNING CLIMATE
Structure school goals Communicates school goals	Manages the curriculum. Supervises and assesses instruction. Monitors student progress	Ensures teaching time. Maintains a visible presence in the school. Provides incentives for teachers. Promotes professional development. Encourages learning

Source: Compiled by the author.

- Statistical Analysis

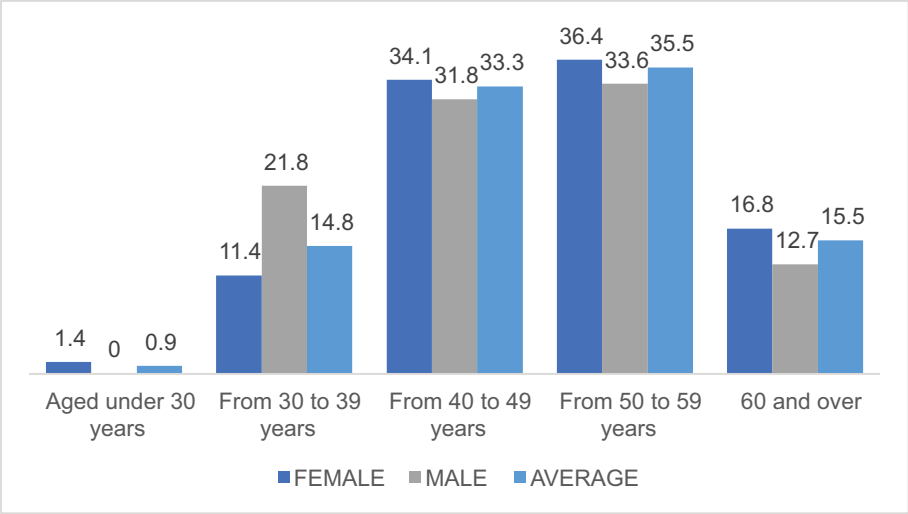
Statistical analysis was carried out using SPSS 28 for Macintosh. The mean scores obtained in each of the categories were calculated to simplify the analysis, as well as bivariate analysis through contingency tables with some questions associated with pedagogical leadership practices compared by sex. Afterwards, the trends in school leadership in Salesian schools in the Americas and the aspects that require improvement in the continent were assessed. It is recognized as a limitation of this study that the analysis responds to the opinions of the participating principals, without a means of control to counteract the possible subjectivity of the responses.

Results

Chart I shows a bivariate analysis with data on the sex and age groups of the participants. Sixty-seven percent of the respondents were women and 33% were men. The age range with the highest percentage is between 50 and 59 years of age (36%), followed by 40 and 49 years of age with 33%.

Received: 07-12-2022 Accepted: 05-05-2023

CHART I. Age Ranges*Gender of Salesian Principals of the Americas (n=330)



Note: Percentages are reported by column. X^2 (4) = 7.9, p<0.01.
Source: Compiled by the author.

We can say that in the population surveyed, female principals are twice as many as male principals in general terms (67%), except in the 30-39 age group, where men (21.8%) have a higher percentage than women (11.4%). This shows that Salesian schools are largely led by women throughout the continent.

Regarding gender distribution by country, the countries with the highest percentage of women in school leadership are Honduras, Nicaragua, Puerto Rico, Dominican Republic, Venezuela, Bolivia, Brazil, El Salvador, Ecuador, and Chile (Table II).

The study assumes that women principals tend to perform more pedagogical leadership practices when these refer to relational aspects, participation, and ability to work with others, compared to men. Of the variables corresponding to pedagogical leadership, the means were calculated for each group of practices or categories and subsequently, questions were selected from the set of variables that most stand out in this type of leadership: supervision of teaching practice and evaluation;

TABLE II. Principals distributed by gender and countries (n=330)

	Sex		Total
	Female	Male	
United States	57.1%	42.9%	100.0%
Argentina	59.0%	41.0%	100.0%
Bolivia	84.6%	15.4%	100.0%
Brazil	77.6%	22.4%	100.0%
Chile	71.4%	28.6%	100.0%
Colombia	52.0%	48.0%	100.0%
Costa Rica	60.0%	40.0%	100.0%
Ecuador	72.7%	27.3%	100.0%
El Salvador	75.0%	25.0%	100.0%
Guatemala	66.7%	33.3%	100.0%
Honduras	100.0%	0.0%	100.0%
Mexico	59.1%	40.9%	100.0%
Nicaragua	100.0%	0.0%	100.0%
Panama	50.0%	50.0%	100.0%
Paraguay	70.0%	30.0%	100.0%
Peru	55.6%	44.4%	100.0%
Puerto Rico	100.0%	0.0%	100.0%
Dominican Republic	91.7%	8.3%	100.0%
Uruguay	64.3%	35.7%	100.0%
Venezuela	90.0%	10.0%	100.0%
Bolivia-EPDB	57.7%	42.3%	100.0%

Note: Bolivia and the Don Bosco Popular Schools (EPDB) are presented separately because the latter corresponds to a set of 300 schools that operate under another management modality and the reading of these data is considered independent for this study.

professional development. conditions for the learning process; collaborative culture; decision making oriented to teaching and learning.

Results by category and gender

As shown in Table III, in a general overview of the set of variables, it is possible to identify tendencies linked more to female leadership,

Received: 07-12-2022 Accepted: 05-05-2023

TABLE III. Principals' Average Pedagogical Leadership by Category and Gender (n=330)

Dimension	Gender		Total
	Female	Male	
Structures school goals	4.40	4.22	4.34
Communicates school goals	3.99	3.85	3.95
Supervises and assesses instruction	4.21	3.94	4.12
Manages the curriculum	4.40	4.19	4.33
Monitors student progress	4.01	3.77	3.93
Ensures teaching time	4.29	4.16	4.24
Maintains a visible presence in school	4.24	3.93	4.13
Provides incentives for teachers	3.96	3.73	3.88
Promotes professional development	4.39	4.12	4.30
Encourages learning	3.90	3.61	3.80

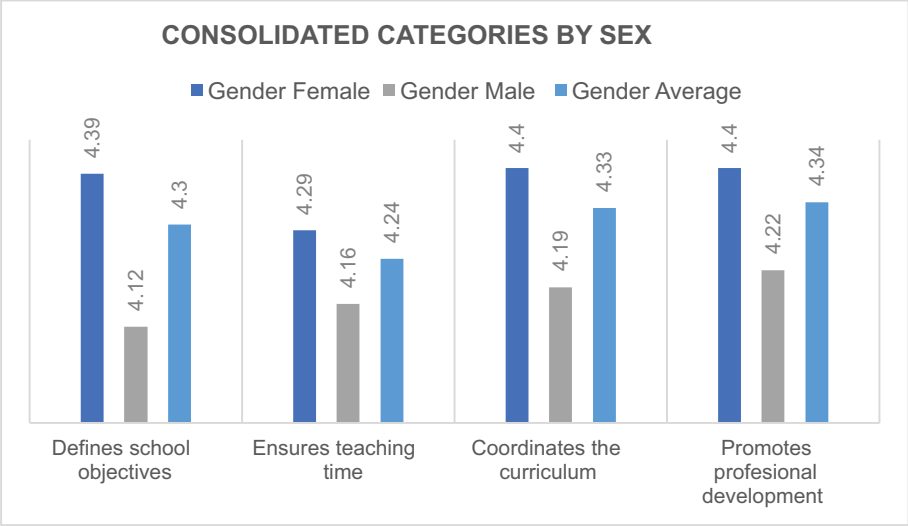
Source: Compiled by the author.

such as supervision and evaluation of teaching, coordinating the curriculum, monitoring student progress, maintaining a visible presence in the school, promoting professional development, and encouraging learning. In the general result, there are consolidated categories and others that are weaker for all principals (Chart II and III).

- Findings on supervision and monitoring of teaching practice

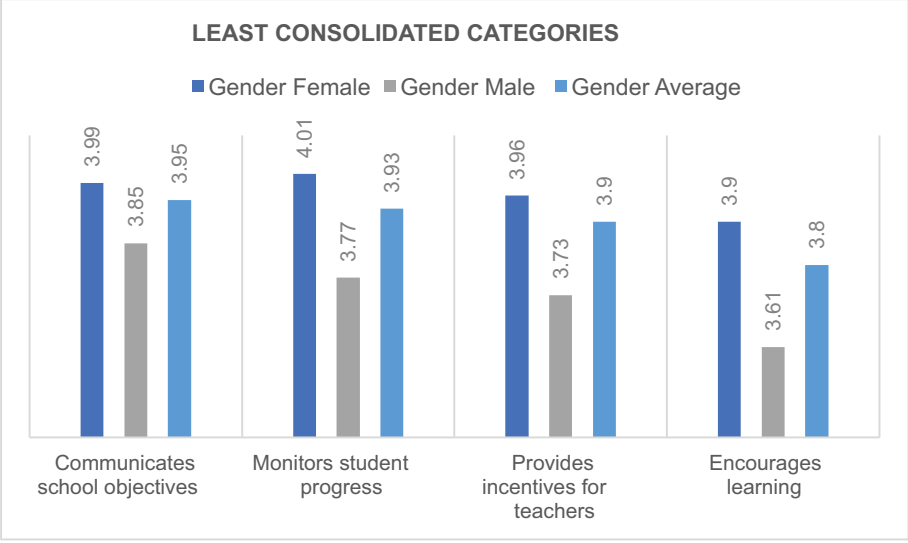
In this set of variables related to classroom teaching practice, the results in Table IV show that female principals carry out this practice more frequently (75.9%) compared to male principals (60.9%). Since these are contingency tables with categorical variables, the Hypothesis of X^2. Thus, it is concluded that H_0 is rejected and there is 95% confidence that there is a relationship between supervision and assessment of teaching practice and sex (X^2 (4) = 11.972, $p<0.05$). That is, we can affirm that there is a relationship between the gender of the principals with respect to the practice of frequent and informal observation in the classroom; women would be showing with greater evidence the importance they give to the process of supervision and teacher accompaniment.

CHART II. Consolidated categories by gender (n=330)



Source: Elaborated by the author.

CHART III. Less established categories



Source: Compiled by the author.

Received: 07-12-2022 Accepted: 05-05-2023

TABLE IV. Reference to Informal Classroom Observations by Gender of the Principal (n=330)

			Sex		Total
			Female	Male	
Conducts informal observations in the classroom on a regular basis	Almost never	Count	5	6	11
		% within sex	2.3%	5.5%	3.3%
	Seldom	Count	9	9	18
		% within sex	4.1%	8.2%	5.5%
	Sometimes	Count	39	28	67
		% within sex	17.7%	25.5%	20.3%
	Frequently	Count	88	44	132
		% within sex	40.0%	40.0%	40.0%
	Almost always	Count	79	23	102
		% within sex	35.9%	20.9%	30.9%
Total		Count	220	110	330
		% within sex	100.0%	100.0%	100.0%

Note: Percentages are reported by column. X^2 (4) = 11.972, p < 0.01.

Overall, Salesian schools in the Americas report this practice as frequent and important (70.9%), however, as a whole, it may be an area for improvement.

Regarding the feedback of the same process, Table IV shows that there is a relationship between the frequency of this practice and the sex variable, being a practice performed more frequently by women (82.3%) than by men (69.1%). The Hypothesis test rejects H_0 and confirms this relationship with 95% confidence (X^2 (4) = 13.781, p < 0.01), i.e., there is evidence to infer that, in the Salesian school population, positive feedback is associated with the sex of the principal. In the overall result, 77.8% of the principals systematically carry out this practice.

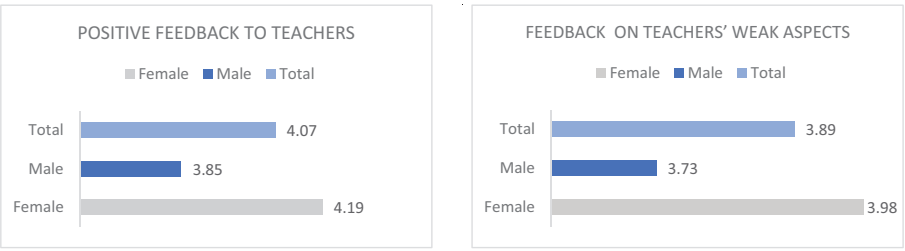
Regarding the feedback of weaknesses in teachers' classroom performance, it is observed that female principals (76%) perform this practice more frequently than male principals (68.2%). However, the null hypothesis does not allow us to affirm a significant gender association at 95%

confidence with respect to reporting weaknesses in classroom practice to teachers (X^2 (4) = 8.917, p>0.05).

At the general level, 73.3% of Salesian school principals carry out this practice. It is of interest to verify which systems are used and which could be implemented to consolidate the feedback from principals to teachers, particularly in the weak aspects.

The following graphs (IV-V) show the averages between both genders for positive feedback and for weaknesses after classroom observation.

CHART IV- V. Feedback to teachers by Principals' gender (n=330)



Compliled by the author.

Professional Development Results

Enhancing the professional development of teachers and educators in the school community is a great challenge and a pressing need, especially in the wake of the COVID-19 pandemic. Therefore, an important issue is how to promote this dimension to support the acquisition of new skills and tools, in addition to generating a culture of collaborative work.

In response to the question about ensuring that the training carried out by school personnel is consistent with the objectives of the school, it is possible to determine that there is no statistical difference between men and women in this practice, i.e., the implementation of training in Salesian schools in America, is a concern that is not related to the gender of the principals and that is quite consolidated (90.3%), which indicates that there is a focus on the importance of teacher training and that this is related to the objectives of the school (X^2 (4)=6.664, p>0.05).

Received: 07-12-2022 Accepted: 05-05-2023

One aspect that is linked to collaborative culture and the ability to work with others is the achievement of participation of all staff in important school activities. The results show a significant association with the gender of the principals, as shown in Table V, where promoting the participation of all staff in activities is more associated with female principals (92.7%) than male principals (87.2%). Women would be evidencing greater frequency in this practice, at 95% confidence, which is confirmed in the Hypothesis test, although with low intensity (p=0.038).

TABLE V. Staff Participation in the Activities and Incidence by gender of the Principals (n=330)

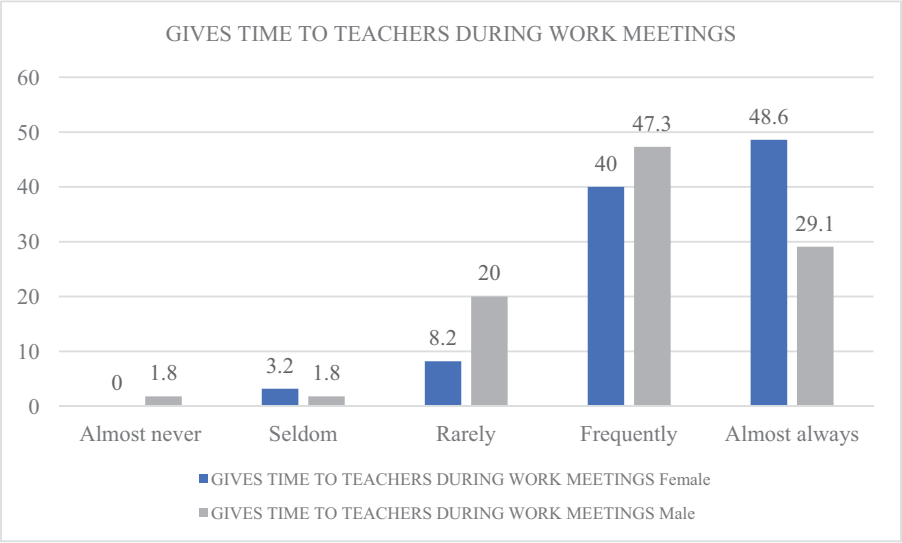
			Sex		Total
			Female	Male	
Achieves participation in relevant activities	Rarely	Count	3	2	5
		% within sex	1.4%	1.8%	1.5%
	Sometimes	Count	13	12	25
		% within sex	5.9%	10.9%	7.6%
	Frequently	Count	95	59	154
		% within sex	43.2%	53.6%	46.7%
	Almost always	Count	109	37	146
		% within sex	49.5%	33.6%	44.2%
Total		Count	220	110	330
		% within sex	100.0%	100.0%	100.0%

Note: Percentages are reported by column. X^2 (3) = 8.433, p< 0.01.

When asked about the time given in work meetings for teachers to share ideas or information about the activities they are carrying out, female principals lead in the frequency of this practice with 88.6% compared to males (76.4%). The hypothesis test allows us to conclude, however, that there is no relationship between these variables at a 95% confidence level. (X^2 (3,261) = p >,05).

Another important aspect in the area of Professional Development is to actively support the use in the classroom of skills acquired during the training offered by the school. The results show that 90.6% of the

CHART VI. Time in meetings for teacher collaborative work associated with gender of principals (n=330)



Elaborated by the author. $X^2(4) = 20.550$, p= 0,000.

principals of Salesian schools in the Americas frequently carry out this practice, which reflects an important consolidation of this attribute linked to the Professional Development of teachers. In turn, the Chi-square test reveals that there is a significant association at 95% confidence, in that this practice is performed more frequently by women (92.7%) than by men, who only obtain 86.3%. ($X^2(4) = 16,751$, $p<0,01$).

The results show a higher frequency of female principals (78.6%) compared to male principals (66.3%) in terms of providing professional growth opportunities for teachers. The hypothesis test confirms that there is a relationship between this variable and the sex of the principals, at 95% confidence. (X^2 (4) = 15,051, $p<0,05$).

Therefore, regarding Professional Development in Salesian schools in America, the results show similarity with the theory, confirming that female pedagogical leadership manifests important skills that favor the accompaniment of teaching practice and the creation of conditions to achieve greater professional growth.

Received: 07-12-2022 Accepted: 05-05-2023

- Conditions for the learning process

Table VI shows the averages obtained by principals in the category of Ensuring teaching time.

TABLE VI. Ensuring teaching time by gender (n=330)

	Sex		
	Female	**Male**	**Total**
Ensures that class periods are not interrupted.	4.54	4.29	4.45
Ensures that students do not leave the classroom during class periods	4.59	4.51	4.56
Ensures that students who are tardy or absent suffer consequences for missing learning time.	4.46	4.34	4.42
Encourages teachers to use classroom time to teach and practice new skills and concepts.	3.54	3.56	3.55
Tries not to let complementary and/or extracurricular activities interfere with the class period.	4.31	4.11	4.24

Source: Compiled by the author.

In this regard, a weakness observed by the group of principals is related to teaching and practicing new skills and concepts in classroom time (3.55). Teachers need to be accompanied in these processes of teaching innovation, which is also related to the permanent capacity to learn new skills, tools and concepts. Considering the value of classroom time, it is necessary to take advantage of and adequately organize what is done there, and principals can play a key role in this by permanently motivating teachers and creating the conditions for them to put the acquired knowledge into practice.

Collaborative culture and decisions for teaching and learning

The next set of questions, related to collaborative culture and decision making, shows the trends observed in the sample (n=330).

Collaborative culture and decision making oriented to teaching and learning obtain a good evaluation in general terms. The association

between sex of principals and the questions described above shows that there is a statistically significant relationship when these functions are performed by women only in the case of question 2.3 "consults the school's academic objectives with teachers when making curricular decisions." ($X^2(4) = 15.633$, $p<0.01$).

TABLE VII. Collaborative culture and teaching and learning-oriented decisions by principals' gender (n=330)

	SEX		
	FEMALE	MALE	TOTAL
1.3 Uses needs assessment or other formal and informal methods to ensure staff contribution to the development of objectives.	4.2	4.0	4.2
2.2 Discusses the school's academic objectives with teachers in working meetings with them.	4.5	4.2	4.4
2.3 Consults the school's academic goals with teachers when making curricular decisions.	4.3	4.0	4.2
3.2 Considers student learning outcomes to evaluate teaching.	4.5	4.3	4.4
7.2 Visit classrooms to meet and share school issues with teachers and students.	4.4	4.2	4.3

Source: Compiled by the author.

Discussion and conclusions

The results obtained, which represent Salesian schools in the Americas in 21 countries, show important strengths in terms of the ability to define the school's objectives (4.34); coordination of the curriculum (4.33); promoting professional development (4.30) and ensuring teaching time (4.24). All of these are of great importance in the exercise of pedagogical leadership centered on learning.

The set of variables explored to establish pedagogical leadership (Table I) show that, in this group of schools in the Americas, it is possible to observe a growing model of pedagogical leadership focused on learning, whose educational purposes include: setting educational goals, planning the curriculum, evaluating teachers, and teaching and promoting teacher professional development (Hallinger, 2005; Robinson et al., 2009).

Received: 07-12-2022 Accepted: 05-05-2023

Aspects that require attention are those related to encouraging learning (3.80); providing incentives for teachers (3.88); monitoring student progress (3.93) and communicating school objectives (3.95). In other words, the functions linked to the management of the teaching program according to Hallinger, in the case of these results, show strengths in terms of curriculum management and in the supervision and evaluation of teaching; however, it is necessary to strengthen the monitoring of the progress made by students with respect to their learning process.

Consequently, it is necessary to promote a leadership style that distributes responsibilities and empowers intermediate leaders to promote learning. Likewise, understanding the school as a Professional Learning Community and generating spaces for it may constitute one of the great strategies for substantive school improvement, strengthening teacher leadership and their ability to build together learning for teaching (Bolivar, 2012; 2019).

In terms of defining the school's mission, the communicational aspect (3.95) requires improvement, along with those items that favor the development of the learning climate: providing incentives to teachers (3.88) and encouraging learning (3.80), both aspects related to strengthening motivation in teachers and students.

Regarding pedagogical leadership, we see that in this set of schools, strong female leadership is exercised, in which the theory about their affinity is corroborated in various categories and functions, highlighting a significant association when these are performed by women (Carrasco and Barraza, 2021).

Other aspects, such as the visibility of the principal in the school, appear as a greater challenge for men (3.93) than for women (4.24); or the supervision and evaluation of teaching, presents more favorable results for female principals (4.21) than for male principals (3.94).

It is worth noting that female principals have a higher frequency in the various practices evaluated, even though this does not necessarily imply a statistically significant relationship. It could be interpreted "between the lines", that there is a significant empowerment of women in Latin American education, as stated by Cárdenas (2014).

As a general conclusion, this approach to pedagogical leadership has allowed, on the one hand, to characterize the most consolidated dimensions and those that should be promoted more strongly by Salesian schools in America to achieve quality learning, thus opening an opportunity for

exchange between countries and regions to share good practices and leadership experiences among their principals and management teams.

It has also made evident the strong female pedagogical leadership that is being developed in the continent, showing promising results for the development of a leadership focused on learning, which allows valuing and projecting this approach as a shared and complementary experience among principals, opening the possibility for future research to give continuity to these findings. The pedagogical leadership approach finds various links with the Salesian educational proposal, so it is desirable to further reflect on the elements and characteristics that make this type of leadership, a concrete operational form of developing a Salesian leadership in the school.

At a crucial time for making decisions regarding the future of post-pandemic education, the transformation that presses for a profound change, capable of generating a better connection between school practice, social demands, and the needs of students in the society of the 21st century, is mostly in the hands of Latin American women. It is possible to affirm, therefore, that school leadership with a female face constitutes an opportunity: "Think crisis - Think female".

Finally, the practices associated with female leadership constitute a learning opportunity for principals as a whole, especially in terms of the competencies that ensure learning. Therefore, considering the presence of female leadership could further enrich the work of management teams, whether in the management role or in intermediate leadership directly linked to pedagogy.

Bibliographical references

Arroyo, D. (2020). Liderazgo de mujeres en la educación: una perspectiva desde las lideres escolares chilenas. *Revista INTEREDU, 3*(2), 9-32. https://bit.ly/3S73a4f

Barber, M. & Mourshed, M. (2007). *Cómo hicieron los sistemas educativos con mejor desempeño para alcanzar sus objetivos*. https://mck.co/3qW6t1l

Bendikson, L.; Robinson, V. & Hattie, J. (2012). Principal instructional leadership and secondary school performance. *Set: Research Information for Teachers, 1*, 1-8. https://bit.ly/3oDyM3W

Received: 07-12-2022 Accepted: 05-05-2023

Bolívar, A. (2010). ¿Cómo un liderazgo pedagógico y distribuido mejora los logros académicos? Revisión de la investigación y propuesta. *Magis, Revista Internacional de Investigación en Educación, 3* (5), 79-106. https://bit.ly/35WLFvT

Bolívar, A. (2012). *Políticas actuales de mejora y liderazgo educativo.* Archidona, Málaga. Ediciones Aljibe.

Bolívar, A. (2019). *Una dirección escolar con capacidad de liderazgo pedagógico*. Editorial La Muralla.

Bolívar, A., & Murillo, J. (2017). *Mejoramiento y liderazgo en la escuela. Once miradas*. CEDLE; Santiago de Chile, (p. 71-79).

Brescoll, V. (2016). Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders, *The Leadership Quarterly*, 27(3), 415-428.

Bush, T. & Glover, D. (2014). School leadership models: what do we know? *School Leadership y Management, 34*:5, 553-571, DOI: 10.1080/13632434.2014.928680

Cáceres, M., Trujillo. J., Hinojo, F., Aznar, I., & García, M. (2012). Tendencias actuales de género y el liderazgo de la dirección en los diferentes niveles educativos. *Educar, 48(1),* 69-89.

Cárdenas de Sanz, M. (2017). En busca del liderazgo femenino: el recorrido de una investigación. Ed. Bogotá, D. C. Colombia: Universidad de los Andes, 2017.

Carrasco, A. & Barraza , D. (2021). Una aproximación a la caracterización del Liderazgo Femenino. El caso de directoras escolares chilenas. *RMIE, 26*(90), 887-910. https://bit.ly/3cDj329

Carrillo, N. (Ed.). (2017). *Género y poder: ¿por qué no hay mujeres directivas?* Universitat Oberta de Catalunya. https://bit.ly/3OK3S4M

Castro, M., Mosteiro, M., & Porto, A. (2021). El liderazgo educativo desde una perspectiva de género. In M. Santos Rego, M. Lorenzo Moledo, Anaïs Quiroga Carrillo (Eds.). *La educación en Red. Realidades diversas, horizontes comunes,* (p. 803). Universidad de Compostela.

Congregación Salesiana. (1996). *Actas Capítulo General XXIV. Salesianos y Seglares, compartir el espíritu de la misión.* https://bit.ly/3ztBcau

Coughlin, A., & Baird, L. (2013). *Pedagogical leadership*. Ontario: Queen's Printer for Ontario.

Cruz-González, C., Lucena, C., & Domingo, J. (2020). Female principals and leadership identity. A review of the literature. *The International Journal of Organizational Diversity, 20*, 45-58.

Cuevas, M., García, M., & Leulmi, Y. (2014). Mujeres y liderazgo: Controversias en el ámbito educativo. *Journal for Educators, Teachers and Trainers, 5*(3), 79 – 92.

Farnsworth, S. (2015). *Principal learning-centered leadership and faculty trust in the principal*. Brigham Young University.

FMA (Instituto Hijas de María Auxiliadora). (1996). *Actas del Capítulo General XX*. "A ti te las confío".

FMA (Instituto Hijas de María Auxiliadora). (1997). *Promoción de la mujer*. https://bit.ly/3beg4gf

Fullan, M. (2002). *Las fuerzas del cambio. Explorando las profundidades de la reforma educativa*. Ediciones Akal.

García-Garnica, M., & Martínez-Garrido, C. (2019). Dirección escolar y liderazgo en el ámbito Iberoamericano. *Revista Profesorado, 23*(2). https://bit.ly/3oyI1m3

Gartzia, L., Ryan, M., Balluerka, N., & Aritzeta, A. (2012). Think crisis–think female: Further evidence. *European Journal of Work and Organizational Psychology, 21*:4, 603-628, DOI: 10.1080/1359432X.2011.591572

Güezmes, A., Scuro, L., & Bidegain, N. (2022). Igualdad de género y autonomía de las mujeres en el pensamiento de la CEPAL. *El Trimestre Económico, vol. LXXXIX 1*(353), 311-338. DOI: 10.20430/ete.v89i353.1416

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools, 4*, 221–239.

Hallinger, P., & Heck, R. (2010). Leadership for learning: Does collaborative leadership make a difference in school improvement? *Educational Management Administration & Leadership, 38(4)*, 654–678.

Hallinger, P., Wang, W., Chen, C., & Li, D. (2015). *Assessing instructional leadership with the principal instructional management rating scale*. Springer.

Jiménez, A., Garay, S., & Parraguez, P. (2023). Diagnóstico de las competencias de liderazgo escolar en directivos salesianos. *Revista Alteridad, 18*(2), 222-233. https://doi.org/10.17163/alt.v18n2.2023.06

Kaiser, R., & Wallace, W. (2016). Gender bias and substantive differences in ratings of leadership behavior: toward a new narrative. *Psychology*

Journal: Practice and Research, 68(1), 72–98. https://doi.org/10.1037/cpb0000059

Leithwood, K. (2009). *¿Cómo liderar nuestras escuelas? Aportes de la Investigación*. Santiago de Chile: Fundación Chile.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). Successful school leadership what it is and how it influences pupil learning. Nottingham: Research Report 800.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management,* 28*(1)*, 27-42

Lewis, P., & Murphy, R. (2008). New directions in school leadership. *School Leadership & Management, 28*(2), 127–146.

Llorent-Bedmar, V., Cobano, V., & Navarro, M. (2017). Liderazgo pedagógico y dirección escolar en contextos desfavorecidos. *Revista Española de Pedagogía, 75*(268), 541-564.

Loyo, A. (2019). ¿Qué queda hoy de la gran ola de reformas educativas de los años noventa en América Latina? *Gaceta de la Política Nacional de Evaluación Educativa en México, 4*(12). https://bit.ly/3CXdMxJ

Malcorra, S. (2018). *Pasión por el Resultado*. Ed. Paidós.

Martínez-Uribe, U., & Martínez-Chaparro, A. (2012). Aproximación al perfil psicosocial de las mujeres líderes del programa "Familias en Acción" del Municipio de Bello. *Pensando Psicología. 8*(14), 118–129. http://bit.ly/3Hdt1oj

Montecinos, C., & Cortez, M. (2015). "Experiencias de desarrollo y aprendizaje profesional entre pares en Chile: implicaciones para el diseño de una política de desarrollo docente", *Docencia,* 55(mayo) 52-61.

Murillo, F. (2006). Una dirección escolar para el cambio: del liderazgo transformacional al liderazgo distribuido. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol. 4*, (4), 11-24.

Murillo, F. (Coord.). (2007). *Investigación iberoamericana sobre eficacia escolar.* Bogotá: Convenio Andrés Bello.

Murillo, F., & Román, M. (2013). La distribución del tiempo de los directores y de las directoras de escuelas de Educación Primaria en América Latina y su incidencia en el desempeño de los estudiantes. *Revista de Educación, 361*. DOI: https://doi.org/10-4438/1988-592X-RE-2011-361-138

Navarro, J., Vergara, M., & Eljach, M. (2018). Liderazgo femenino en el escenario educativo: un fundamento para posibles intervenciones psicoterapéuticas y sociales. *Revista AVFT, 37*(5), 489-494.

Omar, A., & Davidson, M. (2001), Women in management: A comparative cross-cultural overview. *Cross Cultural Management: An International Journal, 8*(3/4),pp.35-67.https://doi.org/10.1108/13527600110797272

ONU Mujeres (Entidad de las Naciones Unidas para la Igualdad de Género y el Empoderamiento de las Mujeres). (2020). COVID-19 y liderazgo de las mujeres: Para responder con eficacia y reconstruir mejor. http://bit.ly/3VCXws0

Padilla, M. (2018). Opiniones y experiencias en el desempeño de la dirección escolar de las mujeres en Andalucía. *RELIEVE, 14*(1),1-27. https://bit.ly/3OFmliy

Pont, B., Nusche, D., & Moorman, H. (2008). *Improving school leadership*. Paris, OCDE. http://www.oecd.org/edu/schoolleadership

Rivera-Mata, J. (2013). *Liderazgo, mujer y sociedad en América Latina*. Ed. Lima: Universidad del Pacífico, 2013.

Robinson, V. (2011). *Student-centered leadership.* San Francisco: Jossey-Bass.

Robinson, V. (2017). Hacia un fuerte liderazgo centrado en el estudiante: afrontar el reto del cambio. In J. Weinstein (ed.), *Liderazgo educativo en la escuela: Nueve miradas*, 45-80. Santiago, Ediciones Universidad Diego Portales.

Robinson, V., Hohepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why. Best evidence synthesis iteration [BES]*. Wellington.

Sandberg, S., & Scovell, N. (2015). *Vamos Adelante (Lean in). Las mujeres, el trabajo y la voluntad de liderar*. Londres: wh Allen.

Schein, V. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology, 57*(2), 95–100. https://doi.org/10.1037/h0037128

Vaillant, D. (2022). El día después en Latinoamérica: Pensar el liderazgo educativo en escenarios de incertidumbre. In A. Bolívar, G. Muñoz, , J. Weinstein and J. Domingo (Coords.). *Liderazgo Educativo en Tiempos de Crisis. Aprendizajes para la escuela post COVID*. (pp. 59-74). Universidad de Granada.

Wacjman, J. (1996). The Domestic Basis for the Managerial Career. *The Sociological Review, 44*(4), 609–629. https://doi.org/10.1111/j.1467-954X.1996.tb00439.x

Weinstein, J., & Hernández, M. (2014). Policies to improve the leadership of School Principals in Chile: A comparison against other Latin American school systems. *Psicoperspectivas, 13*(3), 52-68. https://bit.ly/3zTATXK

Weinstein, J., Hernández, M., Cuellar, C., & Flessa, J. (2015). *Liderazgo escolar en América Latina y el Caribe. Experiencias innovadoras de formación de directivos escolares en la región*. OREAL/UNESCO Santiago. https://bit.ly/3PJHHgb

Zednik, R. (2020). "A Shaken World Demands Balanced Leadership". Medium, 15 de abril. http://bit.ly/3OZHUfE

Zenger, J., & Folkman, J. (2021). Research: Women Are Better Leaders During a Crisis. http://bit.ly/3P1Y73S

Contact address: Patricia Lorena Parraguez Núñez. Universidad Pontificia de Salamanca, escuela doctorado. P.º de Canalejas, 38, 54, 37001, Salamanca, Spain. E-mail: plparragueznu.chs@upsa.es

Grade retention and its relationship with socioeconomic and educative variables in Spain

Repetición de curso y su relación con variables socioeconómicas y educativas en España

https://doi.org/10.4438/1988-592X-RE-2023-402-600

Susana Nieto-Isidro
https://orcid.org/0000-0002-6751-9398
Universidad de Salamanca

Fernando Martínez-Abad
https://orcid.org/0000-0002-1783-8198
Universidad de Salamanca

Abstract

The high percentage of repeating students in the Spanish educational system is a recurring problem whose diagnosis has usually been addressed from the student level and their personal and socio-demographic characteristics. In this work, a different point of view is adopted, with the aim of looking for variables at the school level and at the regional level (Autonomous Communities) that are related to the grade retention, whether they are of a socio-demographic and economic nature, or educational factors subject to be controlled. To do this, based on the data from the Spanish PISA 2018 sample on grade repetition, the proportion of repeaters per center is used as the dependent variable, which is a quantitative variable that is related to different context variables added at the center level and with different socioeconomic and educational variables aggregated at the Autonomous Community level, obtained from both the PISA data and other national and international institutional sources. The initial descriptive and correlational analysis serves as a basis for obtaining the relevant variables included in a two-level multilevel model, in which the first level is constituted by the centers and the second level is constituted by the Autonomous Communities. The results

Received: 18-01-2023 Accepted: 29-03-2023

point to the importance of four types of variables at the center level: socioeconomic, demographic, related to teaching resources and related to teaching work. At the Autonomous Community level, some indices related to social inequality and quality of life are relevant, as well as some aggregate variables related to resources and teaching work. The results point to the importance of improving the human and material resources assigned to the centers and to modify some aspects of the teaching work and the teacher-student relationship to reduce the proportion of repeat students in Spain.

Keywords: grade retention, PISA, multilevel analysis, educational policy, secondary education.

Resumen

El alto porcentaje de estudiantes repetidores en el sistema educativo español es un problema recurrente cuyo diagnóstico se ha abordado habitualmente desde el nivel del estudiante y sus características personales y socio-demográficas. En este trabajo se adopta un punto de vista diferente, con el objeto de buscar variables a nivel del centro escolar y a nivel regional (Comunidades Autónomas) que estén relacionadas con la repetición de curso, ya sean de carácter socio-demográfico y económico, o factores educativos susceptibles de ser controlados. Para ello, a partir de los datos de la muestra española de PISA 2018 sobre repetición de curso, se utiliza como variable dependiente la proporción de repetidores por centro, que es una variable cuantitativa que se relaciona con diferentes variables de contexto agregadas a nivel de centro y con diferentes variables socioeconómicas y educativas agregadas a nivel de Comunidad Autónoma, procedentes tanto de los datos PISA como de otras fuentes institucionales nacionales e internacionales. El análisis descriptivo y correlacional inicial sirve como base para la obtención de las variables relevantes incluidas posteriormente en un modelo multinivel de dos niveles, en el que el primer nivel lo constituyen los centros y el segundo nivel lo constituyen las Comunidades Autónomas. Los resultados apuntan a nivel de centro a la importancia de variables de cuatro tipos: socioeconómicas, demográficas, relacionadas con los recursos docentes y relacionadas con la labor docente. A nivel de Comunidad Autónoma son relevantes algunos índices relacionados con la desigualdad social y la calidad de vida, así como algunas variables agregadas relacionadas con los recursos y la labor docente. Los resultados apuntan a la importancia de mejorar los recursos humanos y materiales asignados a los centros así como a modificar algunos aspectos de la labor docente y la relación profesor-estudiante para reducir la elevada proporción de estudiantes repetidores en España.

Palabras clave: repetición de curso, PISA, modelos multinivel, política educativa, educación secundaria.

Received: 18-01-2023 Accepted: 29-03-2023

Introduction and objectives

PISA assessments by the Organisation for Economic Cooperation and Development (OECD) are an important source of data at different scales in order to conduct highly complex analyses at multiple levels: from student level to higher levels, for example, school, country or region involved in these assessments. Large-scale studies allow us to analyse the effect of some education policies that differ from country to country or region to region, including socio-economic elements, investment in education, segregation, culture, etc.

In this paper we focus on one data from PISA —the proportion of students who state they have repeated a grade— and how this factor may relate with certain personal, contextual, socio-economic and education policy variables of interest.

Grouping will be at school level, i.e., repeating a grade will not be analysed as a dichotomous variable at student level, rather it will be used as a variable dependent on a quantitative variable formed by the proportion of students who state they have repeated grouped by school. Data are used from centres participating in the Spanish PISA 2018 sample (1,082 centres), each with quantitative data on the proportion of repeating students at that school. That variable will be related with factors provided by PISA at student level (considering the average of personal and contextual factors at each centre) and with the characteristics of the centre and, given how Spanish education is organised, with socioeconomic, contextual and education policy data of each Autonomous Region.

A descriptive and correlational analysis as well as multi-level analysis will be used. The latter is appropriate for studying nested data at various levels such as in the case of PISA, but from a different perspective; on one hand, the proportion of repeating students will be used as a dependent variable, in this case quantitative. On the other hand, instead of using common student and centre levels, centres will be used as the first level and Autonomous Regions as the second.

Grade repetition and PISA

PISA includes information on repeating a grade based on student statements in the context questionnaire so there may be some distortion due

Received: 18-01-2023 Accepted: 29-03-2023

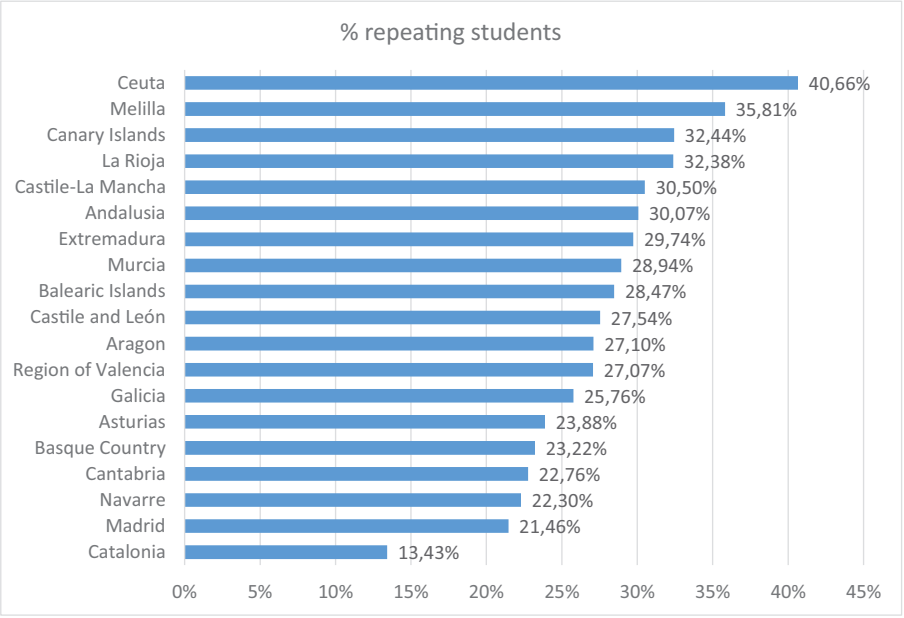
to social desirability (López-Rupérez et al. 2021). Nevertheless, the measure given by PISA can generally be considered a good enough approximation of grade repetition for the purpose of this paper.

An interesting point for this study is that Spain has a very unfavourable ranking in this indicator throughout all waves of PISA that have included this data. In PISA 2018, the percentage ranges from 0.9% in Iceland or 2.5% in the UK to 28.7% in Spain or 30.8% in Belgium; the OECD average is 12.6%.

The percentage of repeating students also ranges widely in the different Autonomous Regions in Spain as shown in figure I for PISA 2018 data.

As seen in figure I, there are major differences between the Autonomous Regions of northern Spain (especially Asturias, Catalonia, Navarre and the Basque Country) and those in the south (especially Andalusia) and the islands (Canary and Balearic Islands). This has already been highlighted in various prior studies (Gil-Flores, 2014; López-Rupérez

FIGURE I. Percentage of repeating students in the Autonomous Regions, PISA 2018



Source: OECD.

Received: 18-01-2023 Accepted: 29-03-2023

et al., 2021) and has been related with economic factors such as the relative poverty of regions (Daniele, 2021) or GDP per capita (Donato and Ferrer, 2012). These differences are also reflected in the regional distribution of school drop-out rates (Calero et al., 2010) as well as primary education results (González-Betancor and López-Puig, 2021). However, the study by Elosúa (2013) with PISA 2009 reading comprehension data finds that differences are due to individual socioeconomic level rather than regional; Donato and Ferrer (2012) also find no association between PISA 2009 science results and regional wealth, and the multi-level model of Rodríguez-Mantilla et al. (2018) for PISA 2015 science performance in Spain does not pinpoint any variable of influence at regional level.

PISA data can also be used to analyse the evolution of grade repetition in Spain. The difference between the various Autonomous Regions generally remains stable over time, although indicators have improved as highlighted by López-Rupérez et al (2021) in a study with PISA 2012, 2015 and 2018 data, as well as García-Perales and Jiménez (2018) with PISA data from 2000 to 2015.

Grade repetition, academic performance and drop-out rates

The debate on the alleged benefits of grade repetition has specific characteristics that make it complicated to obtain a simple answer based on statistical analysis. According to several meta-analysis and literature review studies (Allen et al., 2009; Goos et al., 2021; Jimerson, 2001; Tingle et al., 2012; Valbuena et al., 2021), even though repetition is commonly justified due to poor overall academic performance, it does not appear to have a positive effect on performance and in some cases even shows negative consequences with different effect sizes. Results also differ when comparing repeating students with students of the same age (in different years) or with students from the same year (and who are younger).

One result that seems to be confirmed by available studies is that repetition may have a positive short-term effect but that effect disappears in the long term. In a longitudinal study with secondary students in Belgium, Lamote et al. (2014) show that there is a positive effect in language performance during the year a student repeats but they find a negative effect on that performance in the long term; however, they find no negative effects on self-concept. Similarly, in a longitudinal

Received: 18-01-2023 Accepted: 29-03-2023

study with secondary students in the Spanish Canary Islands, Rodríguez (2022) concludes that repeating has no positive effect on academic performance or student motivational variables such as self-concept. Ehmke et al. (2010) also find that repetition has a positive effect on mathematical performance with secondary students in Germany, and improved mathematical self-concept in repeating students one year after they have repeated a year.

In the case of PISA data, at student level and for all countries, studies on grade repetition seems to concur that there is a negative relationship with academic performance in mathematics, reading and science; this negative relationship between grade repetition and performance in PISA is also maintained in more detailed analyses on some specific groups of students. For example, Hermann and Kopasz (2021) study the relationship between national education policies and the gender gap in mathematics, reading and science with PISA 2012 data and find that grade repetition (grade retention) is the factor with the most consistent correlations with the gender gap: boys obtain better results than girls in countries with a high proportion of repeating students regardless of performance level. The grade level factor - a measure of grade repetition and school starting age— is also a relevant predictive factor for high and low performance in mathematics and science in a paper by Gilleece et al. (2010) with PISA 2006 data in Ireland.

In the specific case of Spain, this relationship has been highlighted by numerous studies with PISA data from different waves and using different approaches and analysis techniques (García-Perales and Jiménez, 2019). For example, when analysing factors associated with high performance in mathematics for the Spanish PISA 2009 sample, Gorostiaga and Rojo-Álvarez (2016) find that grade repetition has a negative effect on performance. Molina et al. (2022) reach the same conclusion with performance data by gender in mathematics for the Spanish PISA 2018 sample; they indicate that it is the variable with the greatest effect on mathematics scores for both genders. Also, García-Pérez et al. (2014) with PISA 2009 data show the negative impact of repetition, especially during primary education or when repeating more than one grade. The same conclusion is reached by Calero et al. (2012) with PISA 2009 in Spain and Italy: grade repetition is a decisive variable for mathematics performance in both countries and particularly in the case of Spain if repeating a grade during primary education or repeating more than once.

Received: 18-01-2023 Accepted: 29-03-2023

One of the characteristics of grade repetition is that it is directly related with school drop-out rates, in other words, with the proportion of individuals who do not complete compulsory studies. Both Jacob and Lefgren (2009) and Hughes et al. (2017, 2018) use longitudinal studies to show the relationship between grade repetition and students dropping out at the age of sixteen.

The same relationship between grade repetition and dropping out is shown in Spain in a study by Lopez-Rupérez et al. (2021), which uses age-appropriateness rate to show that the level of grade repetition at age 15 is a predictor that secondary education students will not graduate. Also, Guio and Choi (2014) use PISA data from 2000 to 2009 to study factors related to school drop-out rates in Spain including grade repetition; the same as the study by Calero et al. (2010) with PISA 2006 data and Choi and Calero (2013) with PISA 2009 data; in all these cases, grade repetition is one of the personal factors associated with dropping out in Spain.

As regards the percentage of students who repeat a grade, international studies show that it varies greatly for the different countries that implement PISA assessments. Goos et al. (2013) indicate that this percentage depends on various factors including beliefs in each society on the advantages and benefits of this practice. In their literature review on the effects of grade repetition policies in the OECD, Valbuena et al. (2021) conclude that this measure is inefficient given the economic, personal and social effects, and they recommend seeking alternatives, as do Calero et al. (2010), Choi and Calero (2013) and García-Pérez et al. (2014). Rodríguez (2022) also warns of this inefficiency due to the detrimental effects on academic performance and different motivational variables such as self-concept; in the same vein as Ehmke et al. (2010) and Tingle et al. (2012).

In the case of Spain, considerable research effort has been dedicated to analysing factors related with grade repetition although most of these studies have focused on the personal and contextual factors of students. For example, Valbuena et al. (2021) show that students from more economically disadvantaged environments and whose parents have less studies are more likely to repeat, the same as López-Rupérez et al. (2021) with data from the Spanish PISA 2018 sample and García-Pérez et al. (2014) with data from the Spanish PISA 2009 sample, which also show the influence of data of birth (being born in the last quarters of the year), gender (boys) and being an immigrant in the probability of students repeating. With data from the Spanish PISA 2019 sample, Cordero et at. (2014) coincide in the

effect of birth month and being an immigrant, along with not attending preschool, family structure or lack of books at home as factors affecting grade repetition. In their study with data from the General Diagnostic Assessment 2009 in Spain, González-Betancor and López-Puig (2016) note that some socioeconomic factors such as the mother's education level and having an unemployed father significantly affect grade repetition.

Arroyo-Resino et al. (2019), with data from PISA 2015 in Spain, relate grade repetition less with socioeconomic factors, which are less important, than with teaching process factors and non-cognitive factors. The most relevant predictors of their model include student aspiration: those with greater aspirations, who expect to complete more than compulsory secondary education, are less likely to repeat.

This study, however, will not focus on students' personal factors which are difficult to change, but rather on the contextual factors of schools, and on socioeconomic and education policy factors at regional level. These are the factors that are easier to change in order to minimise the percentage of repeaters, particularly school-related variables.

Research objectives and questions

The objectives of this study are to analyse the proportion of repeating students in Spanish schools participating in PISA 2018 and relate this proportion with different personal (aggregated by centre), contextual, socioeconomic and education policy data at both centre and Autonomous Region level; and to analyse which of these factors could be controlled by school socio-education policies and teaching to reduce the percentage of repeaters per centre.

The main contributions of this study are: not using repetition as a dichotomous variable at student level but as an aggregated quantitative variable, i.e., by analysing the proportion of repeating students in each school and Autonomous Region. Meanwhile, the analysis will be conducted in two phases: initially with a correlational analysis to find variables related with the proportion of repeaters at the school itself (contextual and educational factors) and at regional level (socioeconomic, development and education policy factors). These variables will then be used for a multi-level analysis with an uncommon approach where the first level will be schools and the second Autonomous Regions.

Received: 18-01-2023 Accepted: 29-03-2023

To achieve these objectives, the following research questions are proposed:

- Q1: What is the current status of grade repetition in schools in Spain and the Autonomous Regions measures by the PISA assessment, and how does it relate with performance?
- Q2: With what centre contextual and education factors can the proportion of repeaters in schools be related?
- Q3: With what regional socioeconomic and education policy factors (Autonomous Region) can the proportion of repeaters in schools be related?
- Q4: Which of the above factors can be changed to reduce the number of repeaters?

This study does not therefore focus on students' personal characteristics associated with grade repetition, but rather on analysing factors associated with grade repetition at meso (school) and macro (system) level.

Method

Design

This *ex post facto* study analysed secondary data. Spanish student and school databases available in the PISA 2018 assessment framework were used[1] (OCDE, 2019, 2020).

Population and sample

The population defined in PISA 2018 in Spain are Spanish compulsory secondary education students aged 15, grouped in their respective schools, which are considered clusters in the sampling process. The OECD implements a two-stage cluster stratified probabilistic sampling. Each Autonomous Region of Spain (AR) is considered a stratum in PISA 2018, which allows a representative sample of both students and schools in each.

[1] https://www.oecd.org/pisa/data/2018database/

The analysis unit in this paper are schools: as explained later, after pre-processing data a final sample was obtained of n=1082 Spanish schools offering compulsory secondary education, distributed by AR as indicated in table I.

TABLE I. Distribution of schools by Autonomous Region

AR	Frequency	Percentage
Andalusia	53	4.9
Aragón	52	4.8
Asturias	55	5.1
Cantabria	55	5.1
Castile and León	59	5.5
Castile-La Mancha	53	4.9
Catalonia	50	4.6
Ceuta	12	1.1
Region of Valencia	52	4.8
Extremadura	54	5.0
Galicia	59	5.5
Balearic Islands	53	4.9
Canary Islands	54	5.0
La Rioja	44	4.1
Madrid	142	13.1
Melilla	8	0.7
Murcia	52	4.8
Navarre	49	4.5
Basque Country	126	11.6
Total	**1082**	**100**

Source: OECD.

Variables and Instruments

In addition to assessing student performance in mathematics, reading and science, PISA tests include context questionnaires given to students, their families, teachers and management teams. These include extensive

Received: 18-01-2023 Accepted: 29-03-2023

information on demographic, socioeconomic, personal, educational, etc., matters of the school environment (OECD, 2019).

The variable criterion of the study was the *proportion of repeaters in the school*. As for predictor variables, based on evidence obtained from previous studies (e.g., Choi and Calero, 2013; Cordero et al., 2014; Goos et al., 2021; López-Rupérez et al., 2021), and taking into account the modules assessed in PISA 2018 context questionnaires (OCDE, 2019), four blocks were included: socioeconomic variables; demographic, school available resources; and teachers' work. All related composite indices available in PISA were included in these blocks. Gender and town size variables, especially relevant in literature, were also included.

It is also important to note that this study is multi-level; with the school being level one and AR level two. Key socioeconomic factors were also obtained from other public databases that are widely used in literature (table II): *per capita income* as a general indicator of AR wealth; spending on education indicators by student and *per capita*; quality of life indicators (such as the Human Development Index); and socioeconomic inequity and inequality indicators (GINI, AROPE and S80/S20 indices).

Other variables included in the study were taken from PISA 2018 databases. Key factors extracted from the student database were added to the schools database, calculating their average. Table III shows the set of variables obtained by block and source database, including the name of each variable in the original databases.

TABLE II. Socioeconomic variables (AR level) included in the study

Variable	**Database**
Per capita income	INE*
Spending on education per capita	INEE**
Spending per student	INEE
Human Development Index.	World Bank Open Data
GINI Index	INE
AROPE (At Risk of Poverty and/or Exclusion) Index	INE
S80/S20 Index	INE

* National Institute of Statistics, ** National Institute for Educational Assessment.
Source: Compiled by authors.

Received: 18-01-2023 Accepted: 29-03-2023

TABLE III. Predictor variables included in the study

Block	Variable	Name in PISA	Database
Socioeconomic	Family socioeconomic level	ESCS	Student
Demographic	Town size	SC001Q01TA	School
Demographic	Ownership	SCHLTYPE	School
Demographic	School size	SCHSIZE	School
Demographic	% of females	ST004D01T	Student
Demographic	Average immigrants	IMMIG	Student
Demographic	Change of school	SCCHANGE	Student
Demographic	Duration of preschool education	DURECEC	Student
School resources	No. of school teachers	TOTAT	School
School resources	Student-teacher ratio	STRATIO	School
School resources	Class size	CLSIZE	School
School resources	ICTs available in the school	ICTSCH	Student
School resources	Computer-student ratio	RATCMP1	School
School resources	Internet Pc-student ratio	RATCMP2	School
School resources	% of teachers with a master's degree	PROAT5AM	School
School resources	% of teachers with a PhD	PROAT6	School
School resources	Shortage of school materials	EDUSHORT	School
School resources	Shortage of human resources	STAFFSHORT	School
Teaching	Student behaviour	STUBEHA	School
Teaching	Teacher behaviour	TEACHEHA	School
Teaching	Disciplinary climate in language	DISCLIMA	Student
Teaching	Direct teacher instructions	DIRINS	Student
Teaching	Perceived feedback	PERFEED	Student
Teaching	Stimulation of reading	STIMREAD	Student
Teaching	Adaptive instruction	ADAPTIVITY	Student
Teaching	Teacher interest	TEACHINT	Student
Teaching	Competitiveness in school	PERCOMP	Student
Teaching	Cooperation in school	PERCOOP	Student
Teaching	Discriminatory school climate	DISCRIM	Student
Teaching	Use of ICTs in school	USESCH	Student
Teaching	Use of ICTs in lessons	ICTCLASS	Student

Source: OECD.

Received: 18-01-2023 Accepted: 29-03-2023

Procedure and data analysis

Before the proposed analyses, available data were pre-processed. As outliers were detected in the higher scores of the *proportion of repeaters* criterion variable data distribution (figure II), in order to avoid possible biases, seven schools were eliminated in which 100% of the students sampled claimed to have repeated a grade. So this study includes 1082 of the 1089 Spanish schools sampled in PISA 2018. Given that the original distribution of schools was altered, the school weighting available in PISA 2018 was not included for data analysis.

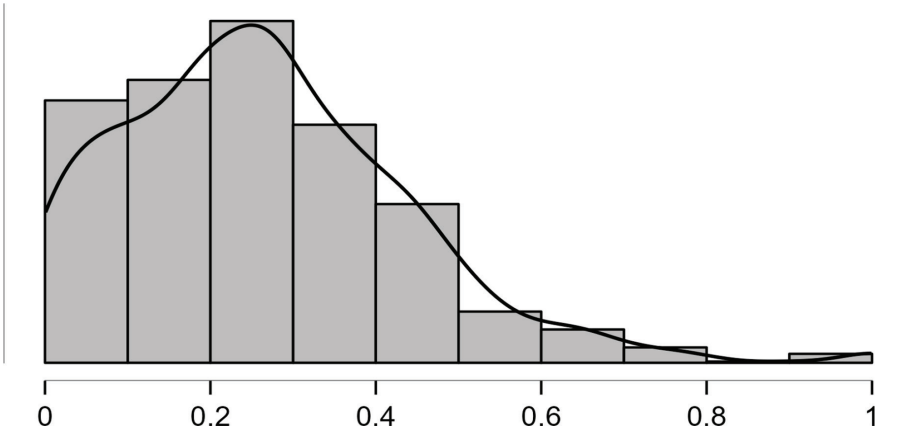
Prior to applying multi-level models to predict the proportion of repeaters we obtained Pearson correlations between the criterion variable and the set of predictor variables, so that initial multi-level models only included variable significantly correlated with the criterion.

Final models were obtained by a step-by-step regression model: the non-significant variable with the lowest weighted parameter was eliminated one by one and the model was repeated until all variables were significant.

A total of three multi-level models were tests in a nested process:

- Model 1: Model including relevant socioeconomic and demographic variables obtained in the correlational analysis.
- Model 2: Based on model 1, it includes socioeconomic and demographic variables as well as those related to school resources.

FIGURE II. Initial distribution of the proportion of repeaters variable



Received: 18-01-2023 Accepted: 29-03-2023

- Model 3: Complete model that includes all significant variables from the four blocks (socioeconomic, demographic, school resources and teaching).

The multi-level models computed included the main effects of the school level and AR level variables as fixed factors and only intercepts of the different ARs as random factors. This decision is based on the exploratory nature of this study, which focused on detecting global factors associated with grade repetition and not detecting how these factors have a differential influence in each AR.

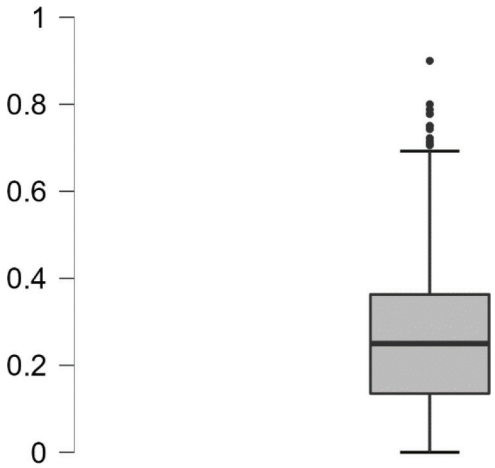
A significance level of 5% was used and analyses were conducted with JASP version 0.16 free software (descriptive and correlational analysis) and with SPSS (multi-level models).

Results

Descriptive and correlational analysis

After eliminating outliers, the proportion of repeaters was distributed in an average value of 25.92% with a standard deviation of 16.54%. Figure III shows how the minimum number of repeating students per school in the

FIGURE III. Distribution of the proportion of repeaters after pre-processing

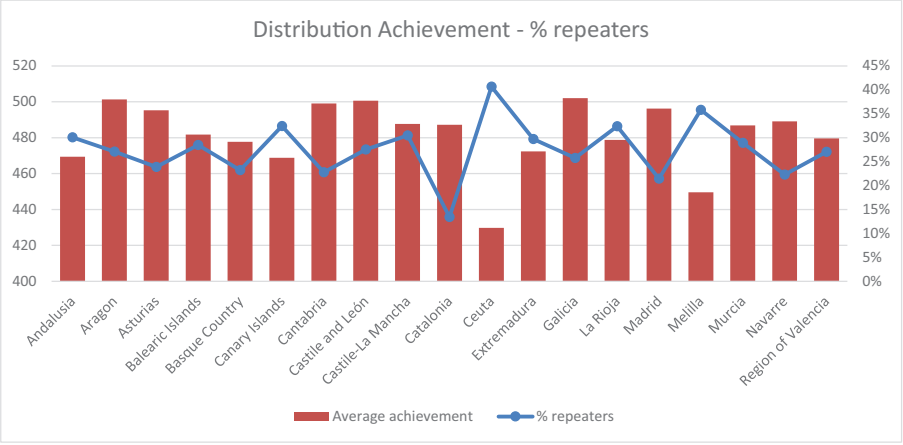


Received: 18-01-2023 Accepted: 29-03-2023

sample is 0% (61 schools), with a maximum of 90% (1 school). The distribution shows a clear positive asymmetry with most schools ranging between 15% and 35% of repeating students.

Also note the high correlation between the proportion of repeaters and average school performance scores in PISA (figure IV). The autonomous regions of Ceuta and Melilla, where the percentage of repeating students is clearly higher than the other ARs, stand out in this sense as they correlate with significantly lower performance levels.

FIGURE IV. Proportion of repeaters by AR and relationship with academic performance



In fact, the correlation between performance assessed in PISA and the proportion of repeaters is very high at both school level and when these variables are aggregated at AR level (table IV).

TABLE IV. Correlations between performance and proportion of repeaters

	School level		AR level	
	r_{xy}	p	r_{xy}	p
Performance in mathematics	-.722	<.001	-.724	<.001
Performance in reading	-.605	<.001	-.662	.002
Performance in science	-.648	<.001	-.675	.002

Received: 18-01-2023 Accepted: 29-03-2023

Regarding correlations between predictor study variables and the proportion of repeating students, table V shows the results in indicators at AR level. Significant correlations can be observed with the overall wealth, quality of life and inequality indicators; direct correlation levels with these latter factors are high.

TABLE V. Correlations between socioeconomic indicators and the proportion of repeaters

	AR level	
	r	p
Per capita income	-.537	.018
Spending on education per capita	-.115	.659
Spending per student	-.180	.489
Human Development Index	-.757	<.001
GINI Index	.596	.007
AROPE Index	.611	.005
S80/S20 Index	.715	<.001

Table VI shows how the proportion of repeaters correlates with other sociodemographic and economic variables in PISA. These indicators are all significant at school level and in terms of AR, only socioeconomic level is relevant in line with the inverse correlation with *Per capita income*. Larger schools in bigger towns which are considered private-charter

TABLE VI. Correlations between sociodemographic factors and proportion of repeaters

	School level		AR level	
	r_{xy}	p	r_{xy}	p
Town size	-.123	<.001	-.311	.195
Ownership	.458	<.001	.291	.227
School size	-.237	<.001	.342	.152
Percentage of females	-.155	<.001	-.169	.488
Average immigrants	.453	<.001	-.326	.172
Change of school	.361	<.001	-.304	.206
Duration of preschool education	-.232	<.001	-.398	.092
Family socioeconomic level	-.715	<.001	-.783	<.001

Received: 18-01-2023 Accepted: 29-03-2023

schools tend to have a lower proportion of repeaters. Meanwhile, schools with a higher percentage of males and a greater rate of immigrant students are associated with higher grade repetition rates. Finally, schools in which students change school more and where students have attended fewer years of preschool are also associated with higher repetition rates.

Many factors associated with school resources are significant at both levels as shown in table VII. At school level, student-teacher ratio (more students per teacher means a lower grade repetition rate, in line with the observations above for school and town size) and management team perception on shortage of material resources at the school (greater needs equal greater repetition rate) stand out. However, the computer-student ratio is not related with the percentage of repeaters, and neither is the level of teachers with higher education. In terms of AR, regions with a lower internet connection rate and greater material and human resource needs are associated with a higher proportion of repeaters.

TABLE VII. Correlations between school resources and proportion of repeaters

	School level		AR level	
	r_{xy}	p	r_{xy}	p
No. teacher	-.105	.001	.158	.518
Student-teacher ratio	-.229	<.001	.421	.073
Class size	-.132	<.001	.345	.149
ICTs available in the school	-.099	.001	-.643	.003
Computer-student ratio	.002	.994	-.136	.580
Internet PC-student	-.037	.245	-.469	.043
Percentage of teachers with master's degree	.076	.033	.197	.419
Percentage of teachers with PhD	.007	.846	-.139	.570
Shortage of school materials	.241	<.001	.602	.006
Shortage of human resources	.164	<.001	.620	.005

Teaching (as perceived by students and management teams) also appears to be significantly associated with the proportion of repeaters, as seen in table VIII. Teaching factors significantly associated with grade repetition at school level include disruptive student and teacher behaviour or negative classroom climate (direct relationship), or frequency of ICT use and cooperative climate (inverse relationship). Teacher skills

Received: 18-01-2023 Accepted: 29-03-2023

TABLE VIII. Correlations between teaching and proportion of repeaters

	School level		AR level	
	r_{xy}	p	r_{xy}	p
Student behaviour	.448	<.001	.389	.100
Teacher behaviour	.189	<.001	.222	.361
Disciplinary climate in language	-.060	.049	.543	.016
Direct teacher instructions	.226	<.001	.675	.002
Perceived feedback	.171	<.001	.505	.027
Stimulation of reading	.037	.228	.601	.007
Adaptive instruction	.039	.195	.399	.090
Teacher interest	-.009	.776	.408	.083
Competitiveness in school	-.062	.042	.287	.234
Cooperation in school	-.299	<.001	-.302	.209
Discriminatory school climate	.243	<.001	.376	.113
Use of ICTs in school	-.105	.001	-.776	<.001
Use of ICTs in lessons	-.173	<.001	-.335	.161

(such as direct instructions and perceived feedback) are directly related with repeating: this association can be interpreted as the teacher being more significantly involved in contexts with greater academic need. At AR level, these teacher skills are once again most significantly associated with the repetition rate. Use of ICTs in school also has an important inverse association at regional level.

Multi-level Models

Firstly, the null model has an intraclass correlation coefficient (ICC) of 8.05%. Despite the literature review (Goldstein, 1995; Snijders and Bosker, 2011) indicating desirable ICC levels of around 10%, the correlational analysis shows numerous significant variables at AR level, which leads us to consider the multi-level analysis to be of interest. Table IX shows parameters of the three final multi-level models obtained. ICC for the complete model is 2.20%, absorbing most intergroup variability with the predictor variables incorporated.

Received: 18-01-2023 Accepted: 29-03-2023

TABLE IX. Multi-level models (L1= level 1, school, L2=level 2, AR)

	Model 1			Model 2			Complete model		
	β	t	p.	β	t	p.	β	t	p.
Intercept	-0.374	-4.08	<.001	1.932	2.76	.011	.988	1.951	.065
GINI (L2)	0.005	2.38	.025	-	-	-	-	-	-
Town Size (L1)	0.008	2.15	.032	0.008	2.16	.031	-	-	-
Ownership (L1)	0.028	4.38	<.001	0.028	4.45	<.001	.021	3.08	.002
Percent. males (L1)	0.154	5.03	<.001	0.154	5.01	<.001	0.139	4.50	<.001
Immigrant rate (L1)	0.085	4.65	<.001	0.083	4.57	<.001	0.095	5.44	<.001
Changes of sch. (L1)	0.154	10.13	<.001	0.154	10.13	<.001	0.144	9.63	<.001
Preschool Ed. (L1)	0.038	2.97	.003	0.037	2.856	.004	0.040	3.21	.001
Socioeconomic lev. (L1)	-0.177	-21.22	<.001	-0.178	-21.32	<.001	-0.158	-19.15	<.001
ICTs available (L2)				-0.059	-2.29	.030	-	-	-
PC-Internet ratio (L2)				-1.787	-2.47	.022	-1.185	-2.34	.029
Student behav. (L1)							0.012	3.56	<.001
ICT use (L2)							-0.147	-5.37	<.001
Direct Instruc. (L1)							0.041	3.73	<.001
Coop. Perception (L1)							-0.026	-2.41	.016

Regarding socioeconomic and demographic variables, of AR indicators (L2), only the GINI index has minor significant effects in model 1. In the final model, school ownership, percentage of males, immigrant student rate, number of school changes and attending preschool are factors at school level (L1) maintained in all models, with direct effects on the proportion of repeaters.

As for school resources, inverse moderate effects can only be observed at AR level (L2) in two variables related with universal access to ICTs in the ARs.

Finally, various factors related to teaching are significant in the complete model. At AR level (L2), frequency of ICT use is a predictor factor of grade repetition. At school level (L1), disruptive student behaviour and greater direct intervention in instructions by the teacher are risk factors associated with higher repetition rates. School environments in which students perceive a cooperative climate, however, are associated with lower grade repetition rates.

Received: 18-01-2023 Accepted: 29-03-2023

Discussion and conclusions

Based on data from the Spanish PISA 2018 sample, grade repetition was analysed by aggregation at school level, calculating the proportion of repeating students at each school, and analysing the relationship between this variable and different indicators at school and regional level. These indicators correspond to four types of variable: socioeconomic, demographic, related to resources and to teaching.

In response to research question Q1, the proportion of repeaters at schools in inversely related with performance in science, mathematics and reading measured by PISA. This result is in line with prior studies that indicate a negative relationship between performance and grade repetition (Calero et al., 2012; García-Perales and Jiménez, 2019; García-Pérez et al., 2014; Gilleece et al., 2010; Gorostiaga and Rojo-Álvarez, 2016; Hermann and Kopasz, 2021; Molina et al., 2022), which questions the use of repetition to improve student performance.

In answer to research question Q2, the school's aggregate socioeconomic level is found to have a major impact on the proportion of repeaters. This relationship between sociodemographic level and repetition coincides with prior research (Calero et al., 2012; Choi y Calero, 2013, Gilleece et al., 2010; González-Betancor and López-Puig, 2016; Goos et al., 2013; Guio and Choi, 2014, López-Rupérez et al., 2021). School demographic variables such as ownership, town and school size, percentage of females and immigrants, and aggregated values of school changes and duration of preschool education of students, are highly significant, in line with studies on individual factors related with grade repetition (Cordero et al., 2014; García-Pérez et al., 2014; González-Betancor and López-Puig, 2016; López-Rupérez et al., 2021; Valbuena et al., 2021). School human and material resources generally have a slight association with the proportion of repeaters at the school, these include teacher-student ratio and class size. Management team perception on shortage of material and human resources is, however, moderately associated with the proportion of repeaters per centre. Finally, student perception of their own behaviour, overall school functioning, perceived climate, and cooperation and teaching, are also associated with the proportion of repeaters.

In response to research question Q3, at Autonomous Region level we can see a direct influence of sociodemographic factors such as greater socioeconomic inequality and lower quality of life, on the proportion

Received: 18-01-2023 Accepted: 29-03-2023

of repeaters, in line with some prior studies (García-Pérez et al., 2014; González-Betancor and López-Puig, 2016; López-Rupérez et al., 2021). Economic inequalities between regions have also proven to have an impact on other aspects of education, such as drop-out rates (Calero et al., 2012) or performance in primary education (González-Betancor and López-Puig, 2021). Daniele (2021) also highlights that the poverty rate in regions of Spain and Italy can predict the rate of students with low performance in mathematics according to PISA, the same as higher levels of inequality, and that regional differences persist even when measures are adjusted based on student ESCS.

However, analysing AR education policy factors like spending on education per student or *per capita*, highlights that even though they are negatively related with the percentages of repeating students in that region, the direct relationships is not significant. Contradictory results can be found on this matter: Donato and Ferrer (2012) find a positive relationship between Autonomous Region spending on education and performance in science in PISA 2009, while Gil-Flores (2014) find no relationship between PISA 2012 performance in mathematics, reading and science and education resources allocated in the Autonomous Regions.

In the multi-level model developed based on the correlational analysis, the complete model includes variables at school level that are sociodemographic (aggregate socioeconomic level by centre), demographic (school ownership, percentage of males, immigrant student rate, number of school changes and attending preschool), and variables related with teaching (disruptive student behaviour, direct intervention in instructions and perception of a cooperative climate). At AR level, the complete model only includes variables related with aggregate teaching resources at schools, such as ICTs available, PC-Internet ratio and use of ICTs.

Lastly, to answer research question Q4 based on multi-level model results, most socioeconomic and demographic variables (such as ownership or aggregate socioeconomic level) stem from the school's context and are difficult to change, although the development of education policies designed to promote preschool education and a more even distribution of immigrant students among schools could be useful. Results also emphasise the importance of improving variables related with school human and material resources (especially universal access to ICTs and their use in the classroom) as well as variables related to teaching, such as creating a cooperative climate and improving student behaviour in the classroom.

In summary, by using a clearly differentiated methodological approach, this study obtains results that are in line with prior studies, thus helping to consolidate the literature review. The results obtained point in two fundamental directions. On the one hand they confirm that fundamental factors associated with grade repetition rate at both school and AR level are socioeconomic and demographic, above factors associated with resources available in schools or teaching. These results must lead us to reflect on the approach and orientation of grade repetition in Spain as they highlight 'a lack of consistency or effectiveness of education compensation mechanisms applied in our country to repeating students' (López-Rupérez et al., 2021, p. 347). On the other hand, taking these inconsistencies into account, our results point toward certain education variables that may have a key effect of education compensation to increase the efficiency of grade repetition. Based on the results obtained, greater analysis is required on how the following key variables impact the effectiveness of grade repetition: adequate use of ICTs in school activities; classroom climate and coexistence at school; open cooperation climate among teachers; and direct instruction style and academic support for students by teachers.

Limitations and future studies

This study used PISA data which, although an excellent source of data, collects information on variables such as grade repetition from student declarations and data may therefore be biased due to social desirability. In line with most research on grade repetition, this study is cross-cutting and thus limits the possibility of finding causal results. In order to analyse cause-effect relationships between the influential variables detected in this paper, it may be interesting to apply causal inference techniques in future studies so as to attribute causal relationships under certain conditions based on panel data. A detailed analysis of education policies implemented and their effects would also be of undoubted value to understand the best way of addressing a complex phenomenon such as grade repetition. Finally, taking into account the significant implications of the relationships found between some variables in this study, future studies are required with a more specific and detailed analysis of the mechanisms behind each of these associations.

Received: 18-01-2023 Accepted: 29-03-2023

Acknowledgements

Grant PID2021-125775NB-I00 funded by MCIN/AEI/10.13039/501100011033/ and by ERDF A Way of Making Europe.

Bibliographical references

Allen, C.S., Chen, Q., Willson, V.L., & Hughes, J.N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis, 31*(4), 480-499. http://doi.org/10.3102/0162373709352239

Arroyo-Resino, D.; Constante-Amores, I.A. & Asensio-Muñoz, I. (2019). La repetición de curso a debate: un estudio empírico a partir de PISA 2015. *Educación XX1, 22*(2), 69-92. https://doi.org/10.5944/educXX1.22479

Calero, J., Choi, Á., & Waisgrais, S. (2010). Determinantes del riesgo de fracaso escolar en España: una aproximación a través de un análisis logístico multinivel aplicado a PISA-2006. *Revista de Educación, Número Extraordinario 2010*, 225-256. https://www.educacionyfp.gob.es/dam/jcr:53d0ecfb-b7b4-4ce9-b2e3-f9280c672801/re201009-pdf.pdf

Calero, J., Escardíbul, J.O., & Choi, Á. (2012). El fracaso escolar en la Europa mediterránea a través de PISA-2009: radiografía de una realidad latente. *Revista Española de Educación Comparada*, (19), 69-104. https://doi.org/10.5944/reec.19.2012.7578

Choi, Á., & Calero, J. (2013). Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma. *Revista de educación, 362*, 562-593. http://doi.org/10.4438/1988-592X-RE-2013-362-242

Cordero, J.M., Manchón, C., & Simancas, R. (2014). La repetición de curso y sus factores condicionantes en España. *Revista de Educación, 365*, 12-37. http://doi.org/10.4438/1988-592X-RE-2014-365-263

Daniele, V. (2021). Socioeconomic inequality and regional disparities in educational achievement: The role of relative poverty. *Intelligence, 84*, 101515. https://doi.org/10.1016/j.intell.2020.101515

Donato, L., & Ferrer Esteban, G. (2012). Desigualdades territoriales en España e Italia: nuevas evidencias a partir de la evaluación PISA-2009.

Revista Española de Educación Comparada. http://doi.org/10.5944/reec.19.2012.7580

Ehmke, T., Drechsel, B., & Carstensen, C.H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. *Studies in Educational Evaluation, 36*(1-2), 27-35. https://doi.org/10.1016/j.stueduc.2010.10.003

Elosúa, P. (2013). Diferencias individuales y autonómicas en el estatus socioeconómico y cultural como predictores en PISA. *Revista de Educación, 361*, 646-664. http://doi.org/10.4438/1988-592X-RE-2013-361-236

García-Perales, R., & Jiménez, C. (2019). Relación entre repetición de curso, rendimiento académico e igualdad en educación: Las aportaciones de PISA. *Revista Educación, Política y Sociedad*. *4*(1), 84–108. https://revistas.uam.es/reps/article/view/12312

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

Gil-Flores, J. (2014). Factores asociados a la brecha regional del rendimiento español en la evaluación PISA. *Revista de Investigación Educativa, 32*(2), 393-410. http://doi.org/10.6018/rie.32.2.192441

Gilleece, L., Cosgrove, J., & Sofroniou, N. (2010). Equity in mathematics and science outcomes: characteristics associated with high and low achievement on PISA 2006 in Ireland. *International Journal of Science and Mathematics Education, 8*(3), 475–496. https://doi.org/10.1007/s10763-010-9199-2

Goldstein, H. (1995). *Multilevel statistical models*. John Wiley & Sons.

González-Betancor, S.M., & López-Puig, A.J. (2016). Grade retention in primary education is associated with quarter of birth and socioeconomic status. *PloS one, 11*(11), e0166431. http://doi.org/10.1371/journal.pone.0166431

González-Betancor, S.M., & López-Puig, A.J. (2021) Student achievement in primary education: region matters more than school, *Compare: A Journal of Comparative and International Education, 51*(8), 1118-1137, https://doi.org/10.1080/03057925.2020.1716304

Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational Research Review, 34*, 100401. https://doi.org/10.1016/j.edurev.2021.100401

Received: 18-01-2023 Accepted: 29-03-2023

Goos, M., Schreier, B.M., Knipprath, H.M.E., De Fraine, B., Van Damme, J., & Trautwein, U. (2013). How can cross-country differences in the practice of grade retention be explained? A closer look at national educational policy factors. *Comparative Education Review, 57*(1), 54-84. https://doi.org/10.1086/667655

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

Guio, J.M., & Choi, Á. (2014). Evolución del riesgo de fracaso escolar en España durante la década del 2000: Análisis de los resultados de PISA con un modelo logístico de dos niveles. *Estudios sobre educación, 26*, 33-62. https://doi.org/10.15581/004.26.33-62

Hermann, Z., & Kopasz, M. (2021). Educational policies and the gender gap in test scores: a cross-country analysis. *Research Papers in Education, 364*, 461-482. https://doi.org/10.1080/02671522.2019.1678065

Hughes, J.N., Cao, Q., West, S. G., Smith, P.A., & Cerda, C. (2017). Effect of retention in elementary grades on dropping out of school early. *Journal of School Psychology, 65*, 11-27. https://doi.org/10.1016/j.jsp.2017.06.003

Hughes, J.N., West, S.G., Kim, H., & Bauer, S.S. (2018). Effect of early grade retention on school completion: a prospective study. *Journal of educational psychology, 110*(7), 974. http://doi.org/10.1037/edu0000243

Jacob, B.A., & Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics, 1*(3), 33-58. http://doi.org/10.1257/app.1.3.33

Jimerson, S.R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School psychology review, 30*(3), 420-437. https://doi.org/10.1080/02796015.2001.12086124

Lamote, C., Pinxten, M., Van Den Noortgate, W., & Van Damme, J. (2014). Is the cure worse than the disease? A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept. *Educational Studies, 40*(5), 496-514. https://doi.org/10.1080/03055698.2014.936828

López-Rupérez, F., García-García, I., & Expósito-Casas, E. (2021). La repetición de curso y la graduación en Educación Secundaria Obligatoria en España: análisis empíricos y recomendaciones políticas. *Revista de Educación, 394*, 325-393. http://doi.org/10.4438/1988-592X-RE-2021-394-510

Molina, D., Molina, E., Sánchez Pelegrín, J.A., & Contreras, J.M. (2022). Estudio por género del impacto de factores contextuales en el rendimiento matemático del alumnado español en PISA 2018. *Revista Complutense de Educación, 33*(4), 645-656. https://doi.org/10.5209/rced.76428

OCDE (2019). *PISA 2018 Assessment and Analytical Framework* https://doi.org/10.1787/b25efab8-en

OCDE (2020). *PISA 2018 Technical Background* https://doi.org/10.1787/1b045c06-en

Rodríguez, D. (2022). Grade retention, academic performance and motivational variables in Compulsory Secondary Education: a longitudinal study. *Psicothema*. https://doi.org/10.7334/psicothema2021.582

Rodríguez-Mantilla, J.M., Fernández-Díaz, M., & Jover, G. (2018). PISA 2015: Predictores del rendimiento en Ciencias en España. *Revista de Educación, 380*, 75-102. http://doi.org/10.4438/1988-592X-RE-2017-380-37

Snijders, T.A., & Bosker, R.J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

Tingle, L.R., Schoeneberger, J., & Algozzine, B. (2012). Does grade retention make a difference? *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 85*(5), 179-185. http://doi.org/10.1080/00098655.2012.679325

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

Contact address: Susana Nieto-Isidro. Universidad de Salamanca, Instituto Universitario de Ciencias de la Educación. Paseo de Canalejas 169, 37008 Salamanca. E-mail: sni@usal.es

Received: 18-01-2023 Accepted: 29-03-2023

Teachers' perceptions of Reciprocal Peer Observation

Percepciones de los docentes sobre la observación recíproca entre iguales

https://doi.org/10.4438/1988-592X-RE-2023-402-601

Mariona Corcelles-Seuba
https://orcid.org/0000-0001-6771-1251
Universitat Ramon Llull. FPCEE Blanquerna

Esther Miquel
https://orcid.org/0000-0002-6557-2574
Universitat Autònoma de Barcelona

David Duran
https://orcid.org/0000-0002-0640-3834
Universitat Autònoma de Barcelona

Abstract

In the last decade, there has been a growing interest in peer observation as a mechanism for teachers' professional development. Based on a reciprocal peer observation intervention, the aims of this study are: 1) to examine teachers' perceptions of reciprocal peer observation; 2) to analyze differences in perceptions as a function of whether teachers were acting as observers or observees; 3) to investigate how writing the post-observation and final reports affects teachers' perceptions of the observation process; and 4) to determine whether teachers' initial perceptions of resistance towards peer observation were reduced after the process of reciprocal peer observation. To that end, a survey was administered after the intervention to 224 teachers (137 in primary education and 87 in secondary education) from 15 public schools in Spain. Results showed that most of the teachers had a positive perception of reciprocal peer observation, as it had helped them to reflect on their own practices and those of their peers. The analysis of differences between the roles showed that observers tended to value the identification of areas to improve their practice, and observees tended to view,

Received: 01-09-2022 Accepted: 07-03-2023

feedback as an opportunity for teaching improvement. Results show that reciprocal peer observation is an excellent opportunity for professional development when teachers perform both roles. Results also underline the importance of writing reports during RPO to identify teaching improvement goals. In addition, after the intervention, teachers indicated reduced resistance to peer observation, showing that a reciprocal, collaborative model can defeat the negative emotions that often emerge in peer observation. Recommendations are suggested for the implementation of this practice in school settings.

Keywords: reciprocal peer observation of teaching, teacher professional development; peer learning, collaborative model, primary and secondary education.

Resumen

En la última década, ha habido un creciente interés en la observación entre iguales como mecanismo para el desarrollo profesional docente. A partir de una intervención basada en una observación recíproca de la enseñanza, los objetivos del presente estudio son: 1) examinar las percepciones de los docentes sobre la observación recíproca entre iguales; 2) analizar las diferencias que perciben los docentes cuando desempeñan los roles de observador y observado; 3) conocer cómo impacta la escritura del informe de post-observación y final en las percepciones del proceso de observación; y 4) identificar si después de la práctica de observación recíproca los docentes reducen su percepción inicial de resistencias hacia la observación entre iguales. Para ello, al finalizar la intervención se administró una encuesta a 224 docentes (137 de primaria y 87 de secundaria) de 15 escuelas públicas en España. Los resultados mostraron que la gran mayoría de los docentes tuvieron una percepción positiva de la observación recíproca ya que les ayudó a reflexionar sobre sus prácticas y las de sus compañeros. El análisis de las diferencias entre los roles de observador y de observado mostró que en el rol de observador los docentes valoraron identificar áreas para mejorar su práctica, y en el de observados, el feedback como oportunidad de mejora, evidenciando que la observación recíproca es una excelente oportunidad para el desarrollo profesional docente en ambos roles. Los resultados también subrayan la importancia de escribir informes para identificar objetivos de mejora de la práctica docente. Además, después de la intervención, los docentes indicaron una reducción de sus resistencias iniciales hacia la observación entre iguales, lo que muestra que un modelo colaborativo recíproco puede disminuir las emociones negativas que a menudo surgen ante esta práctica. Se sugieren recomendaciones para la implementación de esta práctica en la escuela.

Palabras clave: observación recíproca entre iguales, desarrollo profesional docente, aprendizaje entre iguales, modelo colaborativo, educación primaria y secundaria.

Introduction

Contemporary teachers increasingly need lifelong learning to face the growing changes and challenges of 21st century education (OECD, 2020). The latest TALIS report indicates that teachers believe that collaborative forms of professional development, as peer observation, have the greatest impact on their teaching practices. However, only 9% of teachers report providing observation-based feedback to colleagues at least once a month (OECD, 2020). In Spain, the percentage decreases to 5 % (González, 2020). Despite the current need to expand this kind of professional development in the school context, research into the collaborative reciprocal model of peer observation, in which teachers perform both roles (observer and observee), is still relatively underdeveloped, as it is, unfortunately, its practice in the schools. The focus of this study is to analyze in-service teachers' perceptions of reciprocal peer observation as a mechanism for their collaborative professional development.

Teachers' Collaborative Professional Development through Reciprocal Peer Observation

Reciprocal Peer Observation (RPO) involves a pair of teachers with similar degrees of experience and status who mutually agree to observe one or more pedagogical aspects of one another's practice. They use instruments such as an observational grid and/or video recordings to collect evidence of their teaching practices and to offer mutual and constructive feedback with the final goal of improving the teaching practice of both teachers (Corcelles-Seuba et al., 2023; Duran et al., 2020; O'Leary and Savage, 2020).

It is important to distinguish this reciprocal and symmetrical collaborative model of peer observation from other models of peer observation that are asymmetrical, where the observer is an "expert" or a teacher with higher status, and not reciprocal in which the observer is not in turn observed in his or her classroom (Gosling, 2005; 2014; Fletcher, 2018; O'Leary, 2020).

Results of previous research indicate that RPO is a promising training strategy with numerous benefits for teachers' learning when they perform both roles, observee and observer (Roselló and De la Iglesia, 2021;

Received: 01-09-2022 Accepted: 07-03-2023

Versátegui and González, 2019). Observees can receive constructive feedback to improve self-efficacy, self-reflection and self-confidence in their teaching (Bruce and Ross, 2008; Kohut et al., 2007; Motallebzadeh, et al., 2017; O'Learly and Savage, 2020; Shousha, 2015). Observers can learn new methodologies by watching how a colleague manages a classroom (Hendry and Oliver, 2012; Kohut et al., 2007; Motallebzadeh et al, 2017; Thomson et al, 2015; Tenenberg, 2016). Moreover, RPO can provide institutional learning by fostering the development of professional learning communities of teachers (Darling-Hammondet al., 2017).

However, RPO has some challenges. Firstly, it's important to achieve a respectful and collaborative relationship based on trust and support between teachers (O'Learly and Savage, 2020). Research has shown that peer observation can provoke anxiety or leave teachers feeling threatened or judged by more expert colleagues. This resistance to PO is one of the main obstacles to its implementation (Alam et al., 2020; Cosh, 1999; Gosling, 2005). Another challenge is constructive feedback (Roselló and De la Iglesia, 2021). When there is a lack of constructive discussion about teaching practices, teachers run the risk of becoming overly complacent about their own teaching practices, potentially reinforcing their resistance to change (Hammersley-Fletcher and Orsmond, 2005; Gosling, 2005; Shortland, 2004). Other difficulties are lack of time, teaching overload and/or lack of space for joint interaction (Alam et al., 2020; Motallebzadeh, et al., 2017; Hammersley-Fletcher and Orsmond, 2005;; Versátegui and González, 2019). Institutional support in the form of training prior to PO practices and then throughout the whole PO process is essential to overcome resistances and difficulties (Sider, 2019).

Despite its challenges and potential benefits, research into RPO is still relatively underdeveloped (Ridge and Lavigne, 2020), as is its practice in primary and secondary educational contexts (OECD, 2020).

Firstly, prior research has mostly focused on teacher observation using an asymmetrical and non-reciprocal model (O'Leary and Savage, 2020), mainly, on the university level (Fletcher, 2018; O'Leary and Savage, 2020; Shortland, 2004; Zeng, 2020), but less research has analyzed RPO in the actual school context (Alam et al., 2020; Hamilton, 2013; Motallebzadeh et al., 2017; Lam and Lau, 2008; Ridge and Lavigne, 2020). Secondly, research has primary focused on the benefits of peer observation for the observee, but less attention has been paid in the literature

to the observer role (Tenenberg, 2016). Thirdly, many peer observation protocols include writing to promote teaching reflective practice (Hamilton, 2013: Farrell, 2013). However, little research has been done to analyze the impact of report writing on the peer observation process (Lakshmi, 2014; McGuinness and Gibbons, 2005). Finally, considering resistance to PO is one of the main obstacles for its practice in schools, more research is needed in determining whether some of this resistance can be overcome when a collaborative reciprocal approach of PO is adopted (O'Leary and Savage, 2020).

In light of the need to better understand the potential benefits of RPO, the present study examines teachers' perceptions of the RPO process, differences between the roles of observer and observee, and the impact of writing reports. In addition, changes in teachers' degree of resistance to PO were analyzed. Our research questions were the following:

- How do teachers perceive the RPO process (pre-observation, observation and feedback phases)? What are the overall perceived benefits and difficulties of RPO for Teachers' Professional Development?
- What differences do teachers perceive when they perform the observer and observee roles during the RPO process?
- How do writing reports (post-observation and final report) affect teachers' perceptions of the RPO process?
- Does the practice of RPO reduce teachers' perceptions of resistance to PO?

Methods

Context of intervention: Reciprocal Peer Observation conditions and procedure

According to O'Leary and Savage (2020), the success of peer observation for development purposes is contingent on a planned and intentional pedagogical discussion between teachers, based on evidence from their own teaching practice. If peer observation is not well-planned or

structured, it can become counterproductive, generating feelings of resistance and hostility. Therefore, it was necessary to organize the RPO process and structure teachers' pair interactions so that they could serve as a mechanism for peer learning.

Teachers in this study were asked to voluntarily participate in a RPO practice as part of their training program. It was important to ensure voluntary participation and data confidentiality to promote a secure environment for peer learning (Sider, 2019; O'Leary and Savage, 2020). Participants were asked to choose their partners in accordance with the criteria of symmetry in experience and status. Mutual trust and respect between peers were important to ensure the success of observation for developmental purposes (Gosling, 2005; O'Leary and Savage, 2020). In addition, considering the relevance of constructive feedback for the success of RPO (Hammersley-Fletcher and Orsmond, 2005; O'Leary and Savage, 2020) teachers participated in an initial training session. The training included a presentation of the characteristics of the peer observation process (pre-observation, observation and feedback), and some guidelines and practical activities to help observers to offer constructive feedback (adapted from O'Leary, 2020).

After the training session, the pairs of teachers were asked to complete the following three-phase RPO process, followed by a final individual written reflection (O'Leary, 2020):

- *Pre-observation phase*: Teachers were required to conduct at least one pre-observation session in which each pair of teachers had to agree on:
  - the objectives of the observation, which needed to be clear and relevant for participants and agreed upon by both teachers before the observation (Sider, 2019; O'Leary, 2020) (a).
  - the observation criteria, for which purpose they were provided with a grid and were able to make the necessary adjustments (b).
  - the duration of the observation (at least one session for each teacher was required).
  - data collection (in addition to completing the adapted grid, participants were encouraged to record the sessions and to select clips for feedback) (c).
  - the observer and observee roles, following the guidelines that were offered (adapted from O'Leary, 2020) (d).
  - preservation of confidentiality (e).

Received: 01-09-2022 Accepted: 07-03-2023

- *Observation phase*: a minimum of two observations were required, one for each participant, because each teacher performed both the observer and the observee role. In the observation session, after informing students of the reason for the observation, observers were asked to take notes discretely and respectfully about the class, assisted by the grid and/or the audiovisual recording, if agreed upon, without intervening. At the end, the observees were asked to write a brief report on how the session had gone (the post-observational report).
- *Feedback phase*: at least, two feedback session were required, one for each participant. In the feedback session, following a conversational format, the observer invited the observee to make a self-assessment of the session, based on the post-observational report written at the end of the observed session. Then, the observer presented their observations, supported by the evidence collected during the observation, identifying at least one strength and one action requiring explanation. The observee actively participated in this dialogue, and, together, they were asked to set (a few reachable) specific goals for teaching improvement.
- *Individual written reflection*: Based on the different reflections carried out during the RPO process (the individual one at the end of the observed session and the shared ones in the feedback session), the observee was asked to compile a brief report based on a personal synthesis, the final report.

Participants

The participants were 224 in-service teachers from two networks of schools in Spain. 180 of them from 9 schools in a school network in Navarra (Proeducar Hezigarri program) and 44 of them from 6 schools in a Catalan school network (Xarxa de Competències). Teachers from Cataluña conducted the reciprocal peer observation practice during February to May of 2020, and teachers from Navarra during October 2020 to January 2021. 54 participants were males and 170 females. 137 teachers worked in primary education and 87 in secondary education. All of them took part in the study on a voluntary basis, as part of their

training. All participants received written information about the project and gave their consent to participate according to the ethics compliance procedures (Ramrathan et al., 2017). Participants were trained in the peer observation process, and they implemented RPO at least two times, once performing the role of observee and another performing the role of observer.

Instruments

An ad hoc questionnaire entitled *Reciprocal Peer Observation* (*RPO*) *Online Survey* was designed for the purposes of this study.

The *RPO Online Survey* consists of a) 3 sociodemographic items (gathering data on school, educational stage, and gender); b) 59 items organized into three sections (pre-observation, observation, and feedback phases) based on the PO protocol described by O'Leary (2020) to collect teachers' perceptions of the PO process. From those, 47 were close-ended questions answered using a Likert scale (1-4), ranging from strong disagreement (1) to strong agreement (4), 10 items were close-ended binary questions (yes/no), and 2 items -related to difficulties in the pre-observation and feedback phase- were multiple choice questions with closed answers, although participants had the option to add other responses. Categories for the closed answers were constructed by reviewing the main difficulties identified in the previous literature about the topic.

Finally, 7 close-ended questions answered using a Likert scale (1-4) conformed a scale to evaluate teachers' resistance to PO. This scale was designed by reviewing previous literature research about teachers' resistance to PO and consisted in two items measuring resistance to the observer role and five items to the observee role. The tool was validated by a panel of five experts in the field of peer learning. They were asked to review, comment, and clarify the meaning of the wording for each item. They provided feedback on the appropriateness of each item to ensure that all items were relevant. To validate the final scale a Cronbach's α test was performed showing high internal consistency (Cronbach's $\alpha = .904$).

The survey took approximately 10 to 15 minutes to complete.

Received: 01-09-2022 Accepted: 07-03-2023

Procedure

Data collection

Data were collected in June 2020 (teachers from Xarxa de Competències) and in February 2021(teachers from Proeducar Hezigarri program) through the *RPO Online Survey* to assess teachers' perceptions of RPO. Participants were asked to voluntarily complete the survey after they had finished the reciprocal peer observation practice. Responses to the survey were kept anonymous to maintain participants' confidentiality. The survey was distributed via email and was open to all participants for a period of three weeks. The response rate was 98%.

Data analysis

Variables on a Likert scale were described using frequency, mean, standard deviation and percentage (Stevens, 1946; Knapp, 1990).

The participants' qualitative responses were analyzed using the content analysis method (Prasad, 2008). Bottom-up categories were developed by one researcher and then validated by a second researcher. In case of disagreement, a consensus was reached through discussion. Afterwards, qualitative variables were described using frequency and percentage.

To analyze differences between teachers' perceptions when performing the observer role and the observee role, and to examine the differences between teachers' resistances before and after the peer observation, Kolmogorov-Smirnov test was used to determine if the data was normally distributed or not. Results showed p value was .000 so that a nonparametric Wilcoxon Signed Ranks Test was used for paired samples with data on a Likert scale (1-4). For nominal data (yes/no), the McNemar test was used. Finally, differences between teachers that did written reports and those that did not were tested using a Chi-Square Test for homogeneity, if application conditions were satisfied. Otherwise, Fisher's Exact Test or the Likelihood Ratio Test was used.

The statistical analysis was performed using IBM® SPSS® Statistics v.21. For all statistical tests, a nominal significance level of 5% ($p<.05$) was established.

Results

1. Teachers' perceptions of the Reciprocal Peer Observation process

1.1 How do teachers perceive the RPO process (pre-observation, observation and feedback phases)?

Most of the teachers felt secure during the pre-observation meeting (96.4%) and found to be useful this meeting (96.4%), the observational grid (89.7%) and the feedback meetings (95.5%). Meanwhile, most of the teachers (74.1%) did not find it difficult to act as observers (see Table I).

A majority agreed that the feedback they had received from their partners would be useful in helping them to improve their professional practice (93.8%), and disagreed (79.5%) with the statement that it was difficult to offer constructive, non-judgmental feedback.

TABLE I. Peer observation phases

Pre-observation phase	***1***	***2***	***3***	***4***	***Mean***	***SD***
The pre-observation meetings are a useful part of the peer observation process	0%	3.1%	39.3%	57.6%	3.54	.55
I felt safe and calm during the planning meetings.	0%	3.6%	35.7%	60.7%	3.57	.54
Observation phase	***1***	***2***	***3***	***4***	***Mean***	***SD***
The observation grid was useful.	2.7%	7.6%	41.5%	48.2%	3.35	.73
I limited myself to observing the elements on the grid that we had agreed upon.	4.9 %	12.5%	47.8%	34.8%	3.13	.81
It was hard not to intervene in the classroom and to concentrate only on observing.	44.2%	29.9%	20.1%	5.8%	1.88	.92
Feedback phase	***1***	***2***	***3***	***4***	***Mean***	***SD***
Feedback meetings are a useful part of the peer observation process.	1.3%	3.1%	38.4%	57.1%	3.51	.62
The feedback my gave me has been useful to me in improving my professional practice.	1.8%	4.5%	37.1%	56.7%	3.49	.67
It was hard for me to make constructive, non-judgmental comments on my colleague's work.	39.3%	40.2%	15.2%	5.4%	1.87	.86

Note: Likert Scale 1 (strongly disagree) - 4 (strongly agree); N=224.

Received: 01-09-2022 Accepted: 07-03-2023

1.2 What are the overall perceived benefits and difficulties of RPO for Teachers' Professional Development?

Regarding the benefits of RPO, teachers responded positively all the items related to TPD, recording mean scores for all items of over 3 out of a maximum 4 points (see Table II).

TABLE II. Benefits of Peer Observation

Teacher Professional Development	***1***	***2***	***3***	***4***	***Mean***	***SD***
Improve my observation skills	2.7%	7.1%	46.9%	43.3%	3.31	.72
Improve my motivation as a teacher	5.4%	13.8%	46.9%	33.9%	3.09	.82
Improve my professional self-esteem and self-confidence	6.3%	15.2%	47.8%	30.8%	3.03	.84
Focus on areas for improvement and on beginning to make changes	0.4%	4.9%	43.8%	50.9%	3.45	.61
Be more aware of my colleague's teaching style and what we do and do not have in common	1.8%	5.8%	42.4%	50.0%	3.41	.68
Reflect on my own teaching through an analysis of others' practices in order to get to know myself better	1.8%	3.1%	43.3%	51.8%	3.45	.64

Note: Likert Scale 1 (strongly disagree) - 4 (strongly agree); N=224.

The four items with the highest scores were those inquiring about how RPO can serve as a tool to reflect about their and their peers' teaching practice. Most of the teachers agreed that RPO had enabled them to identify aspects to improve in their own practice and to begin to undertake changes (94.7%). The process had helped them to reflect about their own practice via the analysis of their peer's practice (95.1%), to gain greater awareness of the similarities and differences between their own teaching and that of their peer (92.4%), and to improve their abilities as observers (90.2%). Teachers also agreed that RPO was an effective mechanism to improve their motivation as teachers (80.8%), and to their enhance self-esteem and professional confidence (78.6%).

It is also worth mentioning that 72.3% of the teachers stated that they were planning to continue with RPO in the future.

Regarding feedback benefits (see Table III), most of the teachers agreed that the RPO process had helped them to learn both to offer (90.2%) and

TABLE III. Feedback session benefits

Participating in peer observation has allowed me to...	***1***	***2***	***3***	***4***	***Mean***	***SD***
Learn to offer constructive feedback.	1.8%	8.0%	46.4%	43.8%	3.32	.69
Learn to accept my colleague's feedback.	1.3%	8.9%	40.2%	49.6%	3.38	.70
Receive constructive feedback and thought-provoking questions.	.9%	4.0%	39.3%	55.8%	3.50	.62

Note: Likert Scale 1 (strongly disagree) – 4 (strongly agree); N=224.

accept constructive feedback (89.8%). A majority (95.1%) also agreed that the peer observation process had allowed them to receive constructive feedback and thought-provoking questions.

In regard to difficulties, a minority of the teachers said they had experienced problems during the pre-observation (35.3%) and feedback (13.8%) phase (see Table IV). Difficulties in the pre-observation meeting

TABLE IV. Difficulties in the pre-observation and feedback phase

	Yes	**No**
Did you have any difficulties during the pre-observation phase?	35.3%	64.7%
Difficulties in pre-observation phase	**Freq**	**%**
Organization (schedules, etc.)	49	50%
Observational grid's adjustments (focus of observation)	46	46.9%
Agreeing about observer / observee roles	3	3.1%
	Yes	**No**
Did you have any difficulties during the feedback phase?	13.8%	86.2%
Difficulties in feedback phase	**Freq**	**%**
Avoiding judgments	11	24.4%
Avoiding giving solutions	9	20%
Prioritizing areas to be improved	10	22.2%
Avoiding giving your opinion	6	13.3%
Finding time for feedback	5	11.1%
Peer dialogue	4	8.9%

Received: 01-09-2022 Accepted: 07-03-2023

were mainly organizational, including problems finding time for the RPO (50%) and for making adjustment to the observational grid to focus the observation (46.9%).

Some of the problems in the feedback phase were mainly linked to the task of offering this constructive feedback, as avoiding judgements (24.4%), giving solutions (20%) or opinions (13.3%). Other difficulties were related to prioritize areas for improvement (22.2%) or to find time to carry out feedback sessions (11.1%).

2. What differences do teachers perceive when they perform the observer and observee roles during the RPO process?

For each stage of the peer observation process, this study looked at potential differences between the perceptions teachers reported when they performed the observer role and those they described when they acted as observees. These differences are presented below, first with regard to the observation phase, then in relation to the feedback phase, and, finally, in terms of the identification of goals to improve teaching practice.

2.1. Differences in the role of observer and observee in teachers' perceptions of the observation phase

The results of the Wilcoxon Signed Ranks Test for paired samples showed statistically significant differences between observers and observees in the observation phase (see Table V).

Firstly, observees were more likely than observers to believe that the presence of an observer had altered the regular process of the classroom (-revers item- M = 2.72 ±1.15 observer; M =2.50 ±1.15 observee; *p= 0.002*). On the other hand, interestingly, observers showed a greater tendency than observees to say that performing their role (as observers) has allowed them to identify potential ways to improve their own practice (M =3.4 ± .76 observer; M=3.27± .72 observee; *p= .006*).

Surprisingly, though, both observers and observees reported low levels of stress during the observation session, and teachers in both roles said they felt little awkwardness about recording class sessions or having their sessions recorded (with scores less than 2). For these items, no statistical differences between the two roles were found.

Received: 01-09-2022 Accepted: 07-03-2023

TABLE V. Differences in the roles of observer and observee in teachers' perceptions of the observation phase

Observer and Observee roles in the Observation phase	**Observer**		**Observee**		
	Mean	***SD***	***Mean***	***SD***	***Sig. (2-tailed)***
My presence did not seriously alter the usual functioning of the classroom (observer) The observer's presence did not seriously alter the usual functioning of the classroom (observee)	2.72	1.15	2.50	1.15	.002**
While I was observing, I felt stressed out, nervous (uncomfortable) (observer) While I was being observed, I felt stressed out, nervous (uncomfortable) (observee)	1.59	.71	1.68	.72	.075
Recording parts of the session made me feel uncomfortable (observer) Having my colleague record parts of the session made me feel uncomfortable. (observee)	1.90	.84	1.71	.76	.334
Observing allowed me to identify areas to improve in my own teaching practice (observer) Being observed allowed me to identify areas to improve in my own teaching practice (observee)	3.4	.76	3.27	.72	.006**

Note: Likert Scale 1 (strongly disagree) – 4 (strongly agree); Wilcoxon Single Rank Test; * p < 0.05, ** p< 0.01, *** p < 0.001; N=224.

2.2. Differences in the role of observer and observee in teachers' perceptions of the feedback phase

Meanwhile, significant differences were found between the roles for all the items connected to results the feedback phase (p <.01, see Table VI).

Results indicate that observees were more likely than observers to appreciate being able to use video recordings of their lessons to prepare for feedback sessions (M=2.50 ± .89 observer; M=2.77 ± 1.03 observee) and to report having started the feedback sessions using their own post-observation report (M=2.37 ± 1.05 observer; M=2.52 ± 1.03 observee). Also, observees were more likely to say that their colleagues conducting the observation had been able to identify areas for potential improvement of their teaching (M=2.96 ± .72 observer: M=3.26 ± .71 observee). Finally, observees were more likely than observers to say that they had successfully adapted to their role as observees during the feedback session, and to report they had understood feedback as an aid for reflection (M=3.43 ± .71 observer; M=3.59 ± .58 observee).

Received: 01-09-2022 Accepted: 07-03-2023

TABLE VI. Differences in the role of observer and observee in teachers' perceptions of the feedback phase

Observer and Observee roles	**Observer**		**Observee**		
	Mean	***SD***	***Mean***	***SD***	***Sig. (2-tailed)***
Having video recordings of my partner's session was useful to me in giving feedback. (observer) Having video recordings of my session was useful to me in getting feedback (observee)	2.50	.89	2.77	1.03	.001**
The start of the session was based on the report my colleague complied after being observed. (observer) The start of the session was based on the report I compiled after being observed. (observee)	2.37	1.05	2.52	1.03	.001**
I was able to identify positive aspects of my colleague's classroom performance. (observer) My colleague was able to identify positive aspects of my classroom performance. (observee)	3.61	.53	3.50	.59	.003**
I was able to identify areas of my colleague's classroom performance with room for improvement. (observer) My colleague was able to identify areas of my classroom performance with room for improvement. (observee)	2.96	.72	3.26	.71	<.001***
My colleague fulfilled his/her role as set out in the "feedback session guide", understanding feedback as an opportunity for reflection. (observer) I understood the feedback I received as an opportunity to reflect on and improve my teaching practice. (observee)	3.43	.71	3.59	.58	.001**

Note: Likert Scale 1 (strongly disagree) – 4 (strongly agree). Wilcoxon Single Rank Test; * p < .05, ** p< .01, *** p < .001; N=224.

However, observers scored higher on the item asking if they had been able identify positive aspects of the observees' lessons (M=3.61 ± .53 observer; M=3.50 ± .59 observee).

2.3. Differences in the role of observer and observee in identifying goals to improve educational practice in feedback session.

Regarding the process of identifying goals to improve educational practice during the feedback session (see Table VII), significant differences between roles were found (p= .029).

Received: 01-09-2022 Accepted: 07-03-2023

TABLE VII. Identifying goals to improve educational practice

Identifying goals to improve educational practice	**Yes**	**No**	***Sig. (2-tailed)***
When you acted as an observer, did you and your colleague use the feedback session to specify a way he/she could improve his/her teaching? (observer)	71.0% (n=159)	29.0% (n=65)	.029*
When you acted as an observee, did you and your colleague use the feedback session to specify a way you could improve your teaching? (observee)	64.7% (n=145)	35.3% (n=79)	

Note: McNemar Test; * p < .05, ** p< .01, *** p < .001; N= 224.

More observers (71%) said that they had identified goals to improve their colleagues' teaching, while the percentage of observees agreeing with this statement was significantly lower (64.7%).

3. How do writing reports (post-observation and final report) affect teachers' perceptions of the RPO process?

Firstly, 67% teachers did the post-observation report after being observed by their colleagues and 75.4% wrote the final report (see Table VIII). Of those who completed the observation report, almost all the teachers (90.4% of 115 teachers) agreed that the goal they later identified

TABLE VIII. Post- observation Report

Post-Observation Report	**Yes**	**No**	**N**
After your class session was observed, did you compile a brief report about how the session had gone?	67%	33%	224
If you came up with a specific objective for improvement during the feedback session, did this objective coincide with what you wrote down immediately after the session when you were observed?	90.4%	9.6%	115
Final Report	**Yes**	**No**	**N**
Did you write a final report at the end of the feedback session?	75.4%	24.6%	224
Did the final report include one or more specific goals for improvement?	86.4%	13.6%	169
Did these goals coincide with what you and your partner had agreed upon at the end of the feedback session?	95.9%	4.1%	123

Received: 01-09-2022 Accepted: 07-03-2023

to improve their educational practice in the feedback phase coincided with the goal set out in their post-observation report. In the final report, teachers included goals to improve their own teaching practice (86.4%) and a large majority of the teachers stated that these goals coincided with the ones they had identified in the feedback session (95.9%).

Most of the teachers (94.8%) found post-observation report useful in their preparations for the feedback session, and agreed (92.8%) that the post-observation report had allowed them to identify aspects of their own practice with room for improvement. In addition, most of the teachers (91.8%) that did the final report felt that it was a useful part of the peer observation process (see Table IX).

Meanwhile, there were statistically significant differences between teachers that did post-observation report and teachers that did not (p=.011) (see Table X) in how they viewed the process of identifying goals to improve their practice.

Teachers that did post-observation report were more likely to have identified goals to improve their teaching practice in the feedback session (70.7%) than teachers that did not compile a post-observation report (52.7%) (Fisher Exact Test (p=.011)).

The results also showed significant differences between the groups on the items asking whether the peer observation process had allowed them to focus on improving aspects of their own practice and to initiate changes ($LRT_{(3,224)} = 11.72$, p = .008; Cramer V'= .227), and about the degree to which they had reflected on their own practice in light of the analysis of the practice of other teachers ($LRT_{(3, 224)} = 9.82$, p = .020; Cramer V'= .215) (see Table XI). Teachers that did post-observation report

TABLE IX. Post-observation Report benefits

Post-observation Report*	***1***	***2***	***3***	***4***	***Mean***	***SD***
Compiling a brief report after my class was observed was useful in helping me prepare for the feedback session.	1.3%	3.9%	48.7%	46.1%	3.39	.632
Compiling a brief report helped me to identify areas of my teaching practice that could be improved.	0%	7.2%	47.4%	45.4%	3.38	.619
Final Report**						
Writing a final report is a useful part of the peer observation process.	.6%	7.7%	52.7%	39.1%	3.30	.63

Note: Likert Scale 1 (strongly disagree) – 4 (strongly agree). *N=152; **N=169.

Received: 01-09-2022 Accepted: 07-03-2023

TABLE X. Post-observation report impact on identifying goals to improve teaching practice

		Did you write a post-observation report?		
		Yes	**No**	***Sig. (2-tailed)***
When you acted as an observee, did you set a specific goal for improvement during the feedback session?	*Yes*	106 (70.7%) *z=2.6*	39 (52.7%) *z=-2.6*	.0011*
	No	44 (29.3%) *z=-2.6*	35 (47.3%) *z=2.6*	

Notes: Fisher's Exact Test; z=Adjusted Residual; * p < .05, ** p< .01, *** p < .001; N= 224.

TABLE XI. Post-observation report impact on peer observation process

		Have you written a post-observation report?		
Participating in peer observation has allowed me…		**Yes**	**No**	***Sig. (2-tailed)***
Focus on areas for improvement and start making changes †	*Strongly disagree*	0 (0%)	1 (1.4%)	.008**
	Disagree	5 (3.3%)	6 (8.1%)	
	Agree	58 (38.7%) *z= -2.2*	40 (54.1%) *z= 2.2*	
	Strongly agree	87 (58%) *z= 3*	27 (36.5%) *z= -3*	
Reflect on my own teaching through an analysis of others' practices in order to get to know myself better †	*Strongly disagree*	1 (.7%)	3 (4.1%)	.020*
	Disagree	2 (1.3%) *z= -2.2*	5 (6.8%) *z= 2.2*	
	Agree	62 (41.3%)	35 (47.3%)	
	Strongly agree	85 (56.7%) *z= 2.1*	31(41.9%) *z= -2.1*	
Learn to accept my colleague's feedback ‡	*Strongly disagree*	2 (1.3%)	1 (1.4%)	.012*
	Disagree	9 (6%) *z=-2.2*	11 (14.9%) *z=2.2*	
	Agree	54 (36.0%) *z=-1.8*	36 (48.6%) *z=1.8*	
	Strongly agree	85 (56.7%) *z=3*	26 (35.1%) *z=-3*	

Notes: †Likelihood Ratio Test; ‡ Chi-square test; z= Adjusted Residual; * p < .05, ** p< .01, *** p < .001; N= 224.

Received: 01-09-2022 Accepted: 07-03-2023

were more likely to strongly agree with the statements in the first (58%) and the second of these items (56.7%) than teachers who had not done the post-observation report (36.5% and 41.9%, respectively).

Results also showed differences between the groups in their belief that participating in the peer observation process had allowed them to learn to accept their peers' feedback ($\chi^2_{(3, 224)} = 10.971$, p = .012; Cramer V'= .221) (see Table XI). Teachers who did the post-observation report were more likely to strongly agree with this statement (56.7%) than teachers that did not write the report (35.1%).

Finally, there were some statistical differences between teachers that did the final report and those that did not (see Table XII) in terms of how likely they were to report having identified goals to improve teaching practice during the feedback session (Fisher Exact test, p = .015; Cramer V'= .165). Teachers that did the final report were more likely to have set goals to improve teaching practice in the feedback session (69.2%) than teachers that did not do the report (50.9%).

TABLE XII. Final report impact on identifying goals of improvement and on peer observation process

		Have you written a final report?		
		Yes	**No**	***Sig. (2-tailed)***
When you acted as the observee, did you set a specific goal to improve your teaching? ‡	*Yes*	117 (69.2%) z=2.5	28 (50.9%) z=-2.5	.015*
	No	52 (30.8%) z=-2.5	27 (49.1%) z=2.5	
Focus on areas for improvement and start making changes †	*Strongly disagree*	0 (0.0%)	1 (1.8%)	.035*
	Disagree	8 (4.7%)	3 (5.5%)	
	Agree	67 (39.6%) z=-2.2	31 (56.4%) z=2.2	
	Strongly agree	94 (55.6%) z= 2.5	20 (36.6%) z= -2.5	

Note: ‡ Fisher Test; † Likelihood Ratio Test; z=adjusted residuals; * p < .05, ** p<.01, *** p < .001; N= 224.

Meanwhile, there were also statistically significant differences between the groups in terms of how likely they were to believe that the RPO process had allowed them to focus on areas for potential improvement of their own practice and to start making changes (LRT (3, N = 224) =8.612, p = .035; Cramer V'= .198). Teachers that did the final report were more likely to strongly agree with this item (55.6%) than teachers that did not write the report (36.6%).

4. Does the practice of RPO reduce teachers' perceptions of resistance to PO?

Lastly, this study looked at changes in teachers' perceptions of their resistance to RPO. At the end of the process, the participants were asked to score their perceived levels of resistance to RPO before and after peer observation. Results were analyzed using the Wilcoxon Single Rank Test for paired samples to compare differences.

Results showed that although the teachers in the study did not have a particularly high degree of resistance prior to the process (mean score between 1 and 2), the study nonetheless recorded a statistically significant decrease in all the types of resistance related to the observer and observe roles after the RPO (p< .001 in all items; see Table XIII).

TABLE XIII. Resistance to Peer Observation (PO) Scale

Resistance to Peer Observation (PO) Scale	Before PO		After PO		
Did you feel worried or uncomfortable?	Mean	SD	Mean	SD	Sig. (2-tailed)
Observer role					
Because you were observing a colleague.	1.66	.69	1.48	.58	< *.001*
Because you had to offer feedback to a colleague.	1.91	.76	1.66	.65	< *.001*
Observee role					
Because you were being observed by a colleague.	1.89	.77	1.52	.59	< *.001*
Because you were receiving feedback from a colleague.	1.81	.75	1.54	.62	< *.001*
Because you felt your professionalism was being questioned or that you were being judged by a colleague.	1.81	.78	1.53	.59	< *.001*
Because of the possibility that the observed class session might not go as you had hoped, desired or planned.	2.09	.77	1.72	.66	< *.001*
Because of the potential for the presence of an observer in the classroom to distract you or your students.	1.69	.69	1.50	.59	< *.001*

Note: Likert Scale 1 (strongly disagree) - 4 (strongly agree) Wilcoxon Single Rank Test; * p < .05, ** p< .01, *** p < .001; N=224.

Received: 01-09-2022 Accepted: 07-03-2023

When it came to acting as observers, after the intervention teachers reported lower levels of resistance to the tasks of offering a peer feedback and observing a peer. The resistance to being observed also decreased, as after the intervention teachers were less concerned about the class not working as planned, and they reported feeling less afraid of being observed by a peer, receiving peer feedback and being professionally judged. In addition, at the end of the process teachers were less likely to be worried about the potential for an observer to disrupt their classroom.

Discussion and Conclusions

The overarching goal of this study was to analyze in-service teachers' perceptions of RPO as a mechanism for professional development.

Regarding the first aim, most of the teachers in the study had a positive perception of the three phases of reciprocal peer observation process and found it beneficial for their TPD. RPO help them to reflect on their own teaching practice and that of their peers, to identify teaching improvement goals, to improve their abilities as observers, and to increase their motivation, self-esteem, and professional confidence. Teachers also reported having received constructive feedback and said they had been able to offer constructive feedback, which is an essential skill for improving educational practices (Cosh, 1999; O'Leary and Savage, 2020). These positive results are in line with those of previous studies (Bruce and Ross, 2008; O'Leary and Savage, 2020; Motallebzadeh, et al., 2017; Shousha, 2015; Kohut et al., 2007), and adds further evidence of the importance of implementing RPO as an embedded training strategy to promote teachers' reflective practice and collaborative professional development in schools (OECD, 2020; Hamilton, 2013). Moreover, most of the teachers expressed an intention to continue to do peer observation in the future, which is especially relevant considering the need in Spain to expand the use of this type of collaborative professional development in schools (OECD, 2020).

Despite these positive perceptions, some of participants expressed difficulties related to focusing the observation and to giving constructive nonjudgmental feedback, which represents a major difficulty and concern in PO (Cosh, 1999; O'Leary and Savage, 2020; Roselló and De la Iglesia, 2021) These are the most challenging tasks that are required to develop a successful RPO process. It has been established that focused observation

can be an essential part of promoting teachers' critical reflection and helping them to produce constructive feedback (Fletcher, 2018; Hammersley-Fletcher and Orsmond, 2005; Gosling, 2005; Sider, 2019). Thus, echoing this previous research, the study strongly suggests that previous training should emphasize the need to reach an agreement on a few explicitly observable items, and to discuss the potential difficulties of the observer and observee in order to clarify the roles each will play during observation and feedback phases. Prior training should also provide teachers with the tools for giving and getting constructive feedback, as this contributes to encourage professional development. Finally, concerns about finding time to meet with their peers were also mentioned in previous studies (Motallebzadeh, et al., 2017; Alam et al., 2020; Sousha, 2015; Verástegui and González, 2019), and they reinforce the relevance of institutional support to overcome these organizational difficulties (Sider, 2019).

Regarding the second aim, analysis of differences between the observer and observee roles offers further evidence of how the RPO process is an excellent opportunity for reciprocal peer learning when participants perform both roles. On the one hand, the observees expressed a greater understanding than observers that feedback was an opportunity for reflection and improvement of their own teaching practice. They especially valued having video recordings of their lessons and using the post-observational report to start feedback session. These elements of the process gave observees a unique opportunity to see their own performance, and the value they attached to these experiences confirms the relevance of introducing these instruments as useful tools for analyzing teaching practice (O'Leary and Savage, 2020). Also, it is worth noting that during feedback, observees were more likely than observers to perceive that there had been less discussion of positive aspects and more focus on potential areas for improvement. This difference in perceptions indicates that this practice was an antidote to one of the main risks of peer observation: complacency (Gosling, 2005). On the other hand, results showed that observers were more likely than observees to think that performing their role (as observers) had allowed them to identify areas for potential improvement in their own practice. This notable finding confirms the results of previous studies that have highlighted the benefits teachers can gain by learning through observation (Thomson et al., 2015; Tenenberg, 2016; Hendry and Oliver, 2012; Kohut, et al., 2007). Moreover, observers were especially focused on offering constructive feedback, and as such,

Received: 01-09-2022 Accepted: 07-03-2023

they were more likely than observees to perceive themselves as able to identify positive aspects of the observees' practice and to specify goals for improving the observees' teaching.

Regarding the third aim, writing post-observational and final reports helped to set specific goals to improve teaching which underline the powerful role that writing can play as a tool for teachers to reflect on their practice (Farrell, 2013; Lakshmi, 2014; McGuinness and Gibbons, 2005). Post-observational report helped teachers to prepare the feedback session, to accept peer's feedback and to enrich the feedback dialogue. This in turn made observees into more active participants in this process, and, consequently, made the process as a whole more symmetrical. Final report is an effective example of how to reinforce these individual spaces for reflection after the peer observation process. These spaces are important in that they promote teachers' sense of agency, as they are an opportunity for observees themselves to specify their own teaching improvement goals.

Finally, regarding the last aim, the study recorded a decrease in all the types of resistance related to the observer and observee roles. Therefore, it supports that a collaborative model of PO with clear goals and guidelines for observer and observee can help defeat the negative emotions that often emerge in response to the evaluative model of PO with classroom observation (Corcelles-Seuba et al., 2023; O'Leary and Savage, 2020).

Despite these contributions, this study has limitations in terms of potential biases related to self-reporting and sample selection. The results are based on teachers' perceptions after the RPO process and were gathered via a single online questionnaire. Although this method allowed the study to reach a considerable sample of in-service teachers and to contribute to the still limited body of RPO research, further qualitative and longitudinal studies should be done to look not only what teachers perceive, but also at what they do in practice. Future research is necessary to analyze the long-term effects of RPO in teachers' real teaching. These longer studies could include data from more than one classroom observation session for each pair, making it possible to analyze the impact of RPO on changes in teaching practices over time. These data could be contrasted with data obtained from direct observations, especially from the feedback sessions. Another limitation of the study is that the sample was extracted from schools that were already interested in RPO, and teachers participated voluntarily, meaning that the teachers who took part might have

been more likely to have positive reactions to RPO. Despite these limitations, the present study contributes to the existing body of knowledge on PO by showing the relevance of RPO as a professional development mechanism. Therefore, it will be helpful to teachers and school leaders who are interested in developing RPO in their schools to promote teacher collaborative professional development for more effective teaching.

Acknowledgments

We would like to thank the teachers and schools that participated in the study.

Funding

This research was funded by MCIN/AEI/ 10.13039/501100011033, Project PID2020-113719RB-100.

Bibliographical References

Alam, J., Aamir, S. M., & Shahzad, S. (2020). Continuous Professional Development of Secondary School Teachers through Peer Observation: Implications for Policy & Practice. *Research Journal of Social Sciences & Economics Review, (1)*, 1, 56-75. https://doi.org/10.36902/rjsser-vol1-iss1-2020(56-75)

Bruce, C. D., & Ross, J. A. (2008). A model for increasing reform implementation and teacher efficacy: Teacher peer coaching in grades 3 and 6 mathematics. *Canadian Journal of Education/Revue canadienne de l'éducation, 31* (2), 346-370.

Cosh, J. (1999). Peer observation: a reflective model *ELT Journal, 53*(1), 22-27. https://doi.org/10.1093/elt/53.1.22

Corcelles-Seuba, M., Duran, D., Flores, M., Miquel, E. & Ribosa, J. (2023). Percepciones docentes sobre observación entre iguales: resistencias, agencia, procedimiento y objetivos de mejora. *Estudios sobre Educación*, *44*, 35-58. https://doi.org/10.15581/004.44.002

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective Teacher Professional Development. Palo Alto, CA: Learning Policy Institute.

Duran. D., Corcelles-Seuba, M., & Miquel, E. (2020). La observación entre iguales como mecanismo de desarrollo profesional docente. La percepción de los participantes de la Xarxa de Competències Bàsiques. *Àmbits de Psicopedagogia i Orientació, 53*, 49-61.

Farrell, T. S. (2013). Teacher self-awareness through journal writing. *Reflective Practice, 14*(4), 465-471. https://doi.org/10.1080/14623943.2013.806300

Fletcher, J. A. (2018). Peer observation of teaching: A practical tool in higher education. *The Journal of Faculty Development, 32*(1), 51-64.

González, A. (2020). *TALIS 2018. Estudio internacional de la enseñanza y el aprendizaje. Informe español. Volumen II*. Ministerio de Educación y Formación Profesional.

Gosling, D. (2005). *Peer Observation of Teaching*. SEDA Paper 118. Birmingham: SEDA.

Gosling, D. (2014). Collaborative Peer-Supported Review of Teaching. In J. Sachs & M. Parsell (Eds.), *Peer Review of Learning and Teaching in Higher Education* (pp. 13–31). Springer Netherlands. https://doi.org/10.1007/978-94-007-7639-5_2

Hamilton, E. (2013). His ideas are in my head: peer-to peer teacher observations as professional development. *Professional Development in Educatio*n, 39 (1), 42-64. https://doi.org/10.1080/19415257.2012.726202

Hammersley-Fletcher, L., & Orsmond, P. (2005). Reflecting on reflective practices within peer observation. *Studies in higher education, 30*(2), 213-224. https://doi.org/10.1080/03075070500043358

Hendry, G. D., & Oliver, G. R. (2012). Seeing is believing: The benefits of peer observation. *Journal of University Teaching & Learning Practice, 9*(1), 7, 1-9. https://doi.org/10.53761/1.9.1.7

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, 39, 121–123.

Kohut, G. F., Burnap, C., & Yon, M. G. (2007). Peer observation of teaching: Perceptions of the observer and the observed. *College teaching, 55*(1), 19-25. https://doi.org/10.3200/CTCH.55.1.19-25

Lakshmi, B. S. (2014). Reflective practice through journal writing and peer observation: A case study. *Turkish Online Journal of Distance Education, 15*(4), 189-204. https://doi.org/10.17718/tojde.21757

Lam, S.F., & Lau, W.S. (2008). Teachers' acceptance of peer coaching: Impact of collegiality and goal orientation. *Journal of School Connections, 1*(1). 3-24.

McGuinness, C., & Gibbons, C. (2005). Peer observation of teaching: Enhancing self-reflection through writing. In D. Gosling (Ed.), *Peer observation of teaching*. (pp. 7-12). London: SEDA (Publication Nº 118).

Motallebzadeh, K., Hosseinnia, M., & Domskey, J. G. (2017). Peer observation: A key factor to improve Iranian EFL teachers' professional development. *Cogent Education,4*(1), 1-12. https://doi.org/10.1080/2331186X.2016.1277456

OECD (2020). *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*, TALIS, OECD Publishing, Paris https://doi.org/10.1787/19cf08df-en

O'Leary, M. (2020). *Classroom Observation. A guide to effective observation of teaching and learning (2n edition).* Routledge.

O'Leary, M., & Savage, S. (2020). Breathing new life into the observation of teaching and learning in higher education: moving from the performative to the informative. *Professional Development in Education, 46*(1), 145-159. https://doi.org/10.1080/19415257.2019.1633386

Prasad, B. D. (2008). Content analysis. *Research methods for social work*, 5, 1-20.

Ramrathan, L., Le Grange, L., & Shawa, L.B. (2017). Ethics in educational research, in L. Ramrathan, L. Le Grange, & P. Higgs (Eds.), *Education studies for initial teacher education* (pp. 432-443). JUTA.

Ridge, B. L., & Lavigne, A. L. (2020). Improving instructional practice through peer observation and feedback. *Education Policy Analysis Archives,* 28(61), 1-29. https://doi.org/10.14507/epaa.28.5023

Rosselló, M. R., & De la Iglesia, B. (2021). El feedback entre iguales y su incidencia en el desarrollo profesional docente. *Revista complutense de educación, 32*(3), 371. http://dx.doi.org/10.5209/rced.70173

Shortland, S. (2004). Peer observation: A tool for staff development or compliance? *Journal of further and higher education, 28*(2), 219-228. https://doi.org/10.1080/0309877042000206778

Shousha, A. I. (2015). Peer observation of teaching and professional development: Teachers' perspectives at the English language institute, King Abdulaziz University. *Arab World English Journal*, *6*(2), 131-1. http://dx.doi.org/10.2139/ssrn.2834383

Received: 01-09-2022 Accepted: 07-03-2023

Sider, S. (2019). Peer coaching in a school in Cairo, Egypt: Implementation, barriers, and pathways to effective adoption. *International Journal of Mentoring and Coaching in Education, 8*(1), 37-51. https://doi.org/10.1108/IJMCE-04-2018-0016

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.

Tenenberg, J. (2016). Learning through observing peers in practice. *Studies in Higher Education, 41*(4), 756–773. https://doi.org/10.1080/03075079.2014.950954

Thomson, K., Bell, A., & Hendry, G. (2015). Peer observation of teaching: the case for learning just by watching. *Higher education research & development, 34*(5), 1060-1062. https://doi.org/10.1080/07294360.2015.1034349

Verástegui, M., & González, P. (2019). Pensadero de maestros: Una dinámica docente de práctica reflexiva. *Profesorado. Revista de Currículum y Formación de Profesorado, 23*(4), 152–173. https://doi.org/10.30827/profesorado.v23i4.11715

Zeng, L. M. (2020). Peer review of teaching in higher education: A systematic review of its impact on the professional development of university teachers from the teaching expertise perspective. *Educational Research Review, 31*, 1-16. https://doi.org/10.1016/j.edurev.2020.10033

Contact address: Mariona Corcelles-Seuba. Universitat Ramon Llull. FPCEE Blanquerna Cister, 34. 08022, Barcelona. E-mail: Marionacs@blanquerna.url.edu

Analysis on access to the Spanish university system and its implications on the Medicine Degree

Análisis sobre el acceso a la universidad española y sus implicaciones en el Grado en Medicina

https://doi.org/10.4438/1988-592X-RE-2023-402-602

Cristina Yanes Cabrera
https://orcid.org/0000-0003-1374-137X
Universidad de Sevilla

Miguel Ángel Ballesteros Moscosio
https://orcid.org/0000-0001-9522-4303
Universidad de Sevilla

Abstract

The university entrance exam should be a mechanism to guarantee equal opportunities for access to higher education. In highly competitive studies, this test is decisive. The purpose of this article is to focus on the implications of the university entrance exam in the Spanish context, specifically on the situation of the Medicine Degree and its subsequent impact on the specialization exams for Resident Medical Interns (MIR). To do this, we will start by analysing the entry point to the system, i.e. access to university studies, and we will analyse with data what is happening in the Spanish context and by Autonomous Community, with cut-off marks, new enrolments and graduates, as well as the response of the Spanish National Healthcare System (SNS) to take in graduates from the university system. In order to develop this work, a longitudinal descriptive approach has been carried out for each of the objectives set, based on the analysis of official documents (documentary analysis technique). The results show that there is great heterogeneity in university entrance qualifications between the different Autonomous Regions, which has a negative effect on equality of opportunities. On the other hand, the average cut-off mark for the Bachelor's Degree in

Medicine increases significantly year after year, and on occasions is close to the maximum. If this trend continues, the system will become saturated. Finally, with regard to the MIR entrance exam, the number of places on offer satisfies the university system, but does not provide a response to the general situation.

Keywords: entrance exam, university, degree in medicine, cut-off mark, Internal Medical Resident.

Resumen

La prueba de acceso a la universidad debería ser un mecanismo que garantizara la igualdad de oportunidades para el acceso a una educación superior. En estudios universitarios de elevado carácter competitivo esta prueba resulta determinante. El propósito de este artículo es centrarnos en las implicaciones de la prueba de acceso a la universidad pública en el contexto español y focalizarlo en la situación del acceso al Grado en Medicina y sus posteriores repercusiones en las pruebas de especialización para Médico Interno Residente (MIR). Para ello, se partirá del análisis de la puerta de entrada al sistema, es decir, del acceso a los estudios universitarios, y se analizarán con datos lo que está sucediendo en el contexto español y por Comunidades Autónomas (CC.AA.), con las notas de corte, las matrículas de nuevo ingreso y los egresados, así como la respuesta del Sistema Nacional de Salud (SNS) para acoger a los egresados del sistema universitario. Para desarrollar este trabajo se ha llevado a cabo un planteamiento descriptivo longitudinal en cada uno de los objetivos planteados, basado en el análisis de documentos oficiales (técnica de análisis documental). Los resultados arrojan que existe una gran heterogeneidad en las notas de acceso a la universidad entre las distintas CC.AA. lo que redunda negativamente en la igualdad de oportunidades. Por otra parte, la media de nota de corte en el Grado de Medicina se incrementa significativamente año tras año, y en ocasiones próxima a la máxima. De seguir esta tendencia se saturará el sistema. Por último, en lo relativo a la prueba de acceso al MIR, la oferta de plazas satisface al sistema universitario, pero no da respuesta a la situación general.

Palabras clave: prueba de acceso, universidad, Grado en Medicina, nota de corte, MIR.

Introduction

University entrance exams are used in many countries around the world, varying in structure and name depending on the place. In general, this test aims to assess, on an equal footing, the level of knowledge and skills of students for access to university. In Spain, the university access exam,

Received: 14-11-2022 Accepted: 21-05-2023

known as "Selectividad", was established in 1974, and has had different names throughout its history, including "Prueba de Aptitud para el Acceso a la Universidad (PAA)", "Prueba de acceso a la Universidad (PAU)" until 2017, and "Evaluación de acceso a la Universidad (EvAU)" and "Evaluación del Bachillerato para el Acceso a la Universidad (EBAU)", since 2018.

The Ministry of Education and Vocational Training, together with the Autonomous Communities, is responsible for the preparation of university entrance examinations. The Ministry establishes the general criteria and guidelines for conducting the tests, and the autonomous communities have the autonomy to organise and manage these tests in their respective territories in terms of the structure of some of their elements and items (Royal Decree 310/2016; Royal Decree-Law 5/2016). In practice, this circumstance seems to result not in a standardised results, but rather in differentiated tests in each of the 17 autonomous communities (Ruiz-Lázaro et al., 2021, p. 235).

The purpose of this paper is to analyse the details of access to university, precisely coinciding with the presentation of a new entrance exam model by the Ministry of Education and Vocational Training, last July, (with fewer tests and a maturity assessment, which is intended to be fully implemented by the 2026-2027academic year). Furthermore, the goal is to structure this analysis around the access to the Bachelor's Degree in Medicine, one of the most sought-after programmes with a high cut-off mark. The study is also justified in the context of the specific social problem of the lack of medical professionals, which has already been highlighted in some studies (Vicenç Martínez Ibáñez et al., 2022) and which has recently been reported in the media and social networks.

All this puts universities in the spotlight in the search for solutions, which to a large extent has been reflected in the Ministry of Universities' measure to increase by up to 15% the places offered in the Degree in Medicine for the next academic year. This measure will be supported by a budget allocation for those universities that decide to implement it. However, some stakeholders in the health sector have raised doubts about the effectiveness of this measure (Echevarría, 2022), arguing that possible solutions should focus on increasing the number of MIR (Medical Internship Residents) places, increasing the number of specialists, or improving working conditions. In order to address this debate, it is advisable to start by analysing the entry point to the system, i.e. access to university studies, and to conduct a data-based analysis on what is happening in the Spanish context regarding cut-off marks, new enrolments

and graduates, as well as the response of the Spanish National Health-care System (SNS) in taking in graduates from the university system.

Status of the issue

The recent doctoral thesis by Judit Ruiz Lázaro is of great interest for an overview of the scientific literature on university entrance examinations in Spain. It provides a general overview of the different types of studies and approaches that have been carried out with respect to the analysis of university entrance in Spain up to 2020 (Ruiz Lázaro, 2021). In the same year when this system of access to higher education studies was launched, the *Revista de Educación* echoed the implications that this procedure could have on the system, and dedicated monographic issue no. 230 to access to higher education. A dozen papers examined the situation extensively, both from a general point of view, including the problem of selection in medical studies (Del Sol, 1974), as well as access systems in seven countries. Since then, and although the body of literature is not large, numerous studies have been published from multiple perspectives, all of them highlighting the complexity of such analysis.

Following the classification proposed by the author in her doctoral thesis on research studies on university entrance in Spain, the first aspect that can be underlined is that no scientific work has been published on "*the entrance model in Spain*" for more than a decade. This category initially includes studies linked to the implementation of the university entrance exam system, followed by an analysis of different aspects of the subject from a more general perspective, focusing on the debate. Apart from the aforementioned monograph in the *Revista de Educación*, the monograph on "Selectividad y Educación" in the journal *Documentación social. Revista de Desarrollo Social*, no. 15, also published in 1974, emphasised at the time what the entrance system could mean in terms of the limited number of places (then numerus *clausus*) implemented in the access to studies. On another level, this aspect, after the various reforms, would lead to what is known as "cut-off marks", a minimum score that determines which students are admitted to a degree course, which depends on the demand and the number of places available at each university, with a great impact on access to higher education studies. Some years later, in 1997, the *Revista de Educación* once again published a monographic

Received: 14-11-2022 Accepted: 21-05-2023

issue, no. 314, on "*University entrance exams*". The studies on the then existing model and its impacts, agreed on two main conclusions: first, that it was necessary to achieve a fair and equitable procedure that, in a valid manner, could ensure access to university studies, and second, that society should realise that if not all students can study what they want, it is not a problem of access but of the mismatch between supply and demand (Muñoz-Repiso and Murillo, 1997, p. 47). On the other hand, "*the autonomy of the Communities and universities cannot and should not, under any circumstances, result in a situation of injustice that violates the principle of equal opportunities for all students*" (Murillo, 1997, p. 62).

In relation to cut-off marks, subsequent regional analyses show that there has been a general progressive increase in these scores (Ruiz-de Gauna and Sarasua, 2011), and that this has a significant effect linked to graduation rates (Jiménez García et al., 2021). In addition, there are several studies showing from different approaches that there are notable differences between test results in the different autonomous communities (Ruiz and González, 1997; Muñoz-Repiso et al., 1997; Boal et al., 2008; Ruiz and González, 2017; Ruiz-Lázaro and González, 2017; Pérez-Cárceles, and Martínez-Martínez, 2019; Mengual, 2019; Ruiz-Lázaro et al., 2021; Faura-Martínez et al., 2022). The differences studied reveal multiple aspects such as disparity in the criteria provided to correctors depending on the university district, or diversity in the structure and contents of various subjects for access to university. This could have socio-economic effects in the different autonomous communities, since the differences in the average access marks determine the ability to study in the region of origin and, consequently, in the case of being below average, be detrimental to the students themselves (Pérez-Cárceles and Martínez-Martínez, 2019, p. 88).

However, the analysis of the difficulties and room for improvement in order to ensure effectiveness of entrance exams undoubtedly requires detailed analyses that take into consideration the multiple variables involved in this process, such as contextual variables related to the results, psychometric characteristics of the tests, alternative access routes to university, or the relevant comparison with the entrance system implemented in other countries, all of which are dimensions addressed by reference publications in Spain (Ruiz Lázaro, 2021, pp. 121-154).

Focusing this context on medical studies, it is worth noting that, globally, access to medical studies in almost all countries of the world is pursued, and has been very competitive for years (Laurence et al., 2013). It

has also become clear that the literature on general selection policies for access to this type of university education is scarce, and that more sophisticated evaluation approaches using multidisciplinary theoretical frameworks are required to address these issues (Patterson, 2018). Admission methods for medical studies in different countries around the world are based on various models of individual assessment (e.g. end-of-school examinations, admission interviews, aptitude tests, etc.). Ideally, the methods should follow general policies. But these models mostly respond to political strategies that are adopted on the basis of both universal and local criteria. Universal criteria are applicable anywhere in the world, and the most common are academic or intellectual capacity and the potential to master professional skills. However, local criteria depend on a country's needs, capacity (e.g. available resources) and socio-cultural values (Soemantri et al., 2020). All of which makes it clear that there is no shortage of elements in the analysis required for the study of university access at this crucial time, to which we hope to make our small contribution.

Background and objectives

Access of future medical professionals to the SNS is subject to two selective tests at two different times. On the one hand, the admission processes for pre-university students to Spanish public universities require them to pass a series of tests structured in two phases, one general and one specific (voluntary), depending on the modality they have previously studied (Organic Law 3/2020). The final mark for admission is the sum of the average mark of the Upper Secondary Education (60%) and that of the selection test (40%) and the result of this mark will determine access to the degree to be studied. Variability in the access mark is, therefore, influenced to a greater extent by the differences in the average mark of the Upper Secondary Education, which is the result of different factors, not least of which is the school where it was studied and other factors specific to the autonomous community, linked, among others, to the different school and assessment culture of each place and the differences in the autonomous community curricula.

On the other hand, access to specialised medical training for graduates in Medicine is determined by means of an annual examination (MIR exam), one single exam for all Spain (Law 44/2003). The test consists of

Received: 14-11-2022 Accepted: 21-05-2023

a multiple-choice exam and reserve questions related to medical studies. The final mark is the sum of the average mark obtained in the Medicine degree -and the doctoral thesis if applicable-, (10%), and the result of the MIR exam (90%). Likewise, the final grade obtained will determine the possibility to choose the speciality in which they wish to train. Both procedures, therefore, place the requirements for the medical profession at a very high level, perhaps one of the most demanding ones in Spain.

If we analyse the different stages of the process, initially, future professionals must access university through 17 different exams, designed by and for each of the Autonomous Communities. This could influence the average cut-off mark required to enter the Medicine degree that in Spanish public universities has been above 13 points (out of a maximum mark of 14) in the last two years. Once admitted to the Spanish university system (SUE), both public and private (private universities establish their own admission criteria, which are not necessarily bound to the entrance grade), students will face a six-year Bachelor's Degree programme in Medicine at a total of 48 Spanish universities (Registro de Universidades Centros y Títulos, RUCT, 2022). In each of the universities, they will study a different curriculum, designed by the different universities, and assessed initially and on an ongoing basis in their verification reports, by the evaluation agencies responsible in each case. Thus, heterogeneity of the processes, until they reach the MIR exam, is a striking aspect. All the more so if students will then have to face a selective test that will define their professional development.

Therefore, although the MIR test is considered to be a determining factor in the entire career path towards the profession, the university entrance mark, as well as the training they receive during their six years at university, will influence their results. Regarding the education received, some studies have shown that, although it accounts for only 10% of the final mark, there is a direct relationship between the marks obtained throughout the medical degree and the score they obtain in the test, objectively showing its influence on the order number they will obtain in the MIR test (Baladrón Romero et al., 2022; Baillès et al., 2020). Concerning the university access mark, a study carried out in the 2008-2014 promotion, in 29 medical schools in 12 Autonomous Communities , shows that there is no correlation between the cut-off mark of each university and performance in the MIR exam (Sentí et al., 2016). In any case, a comprehensive study of the situation would be required, from access to university itself to the possibilities offered by the SNS to university graduates.

To this end, first, we will analyse the situation of the university entrance exam using data that will allow us to find out the degree of participation, in terms of the number of people enrolled, and the average marks for this entrance exam, as well as the situation of the average cut-off mark for the Bachelor's Degree in Medicine. This study will also be addressed in the context of the population range of 17-18 year olds who enter university, as they constitute the majority of the access range. A comparative study of the situation in the different Autonomous Regions will also be carried out, with the aim of determining whether the entrance exam produces the same results based on the Autonomous Region of origin.

In a second stage, we plan to analyse the evolution of the university entrance exam cut-off marks of students of the Degree in Medicine over the last four years, and additionally to do so broken down by Autonomous Community. Likewise, in the academic context and for this period, the enrolment of new students in the different universities will be shown, as well as the number of graduates that the university system provides in these same courses. The aim of this analysis is to assess the advantages and disadvantages of the current system as a starting point for decision-making on the new university entrance exam, which is currently being drawn up.

Third, we present, based on existing data, the possibilities offered by the SNS to take on graduates from the university system, showing how many places, in each of the years under study, are offered for access to the MIR status. The aim in this case is to identify whether or not there is a match between the places offered for the MIR and the number of Medicine graduates, and to try to contribute to the discussion on whether it is necessary to increase the number of places for new entrants to the Medicine Degrees.

Hypothesis and Methodology

For these three objectives of analysis, the following hypotheses are put forward. On the one hand, the system of equal access to university might not be effective. This would determine access to the most sought-after degrees, such as the Bachelor's Degree in Medicine.

On the other hand, the increase in cut-off marks in general, and in particular for the Bachelor's Degree in Medicine, could lead to saturation of the system, if the approach to the university entrance exam is not changed.

Received: 14-11-2022 Accepted: 21-05-2023

Finally, access to the SNS for medical graduates should be rethought. A reasonable solution could be to opt for an alternative university education (degree) that would provide a response for the system.

Design

This is a study based on secondary data analysis (purely quantitative-positivist in nature), using previously collected data and focusing on the objective and quantitative analysis of these data to meet the research objectives.

For its development, a longitudinal descriptive approach has been carried out for each of the objectives set, based on the analysis of official documents (documentary analysis technique) (Bowen, 2009; Peña Vera, 2022) followed by a comparative analysis.

Participants

For the study, we have taken as a reference the 17 and 18 year-olds who sit the entrance exam, as they are the main age range within those who sit the exams. The students who, ordinarily, sit these tests do so in the year in which they turn 18 and, therefore, approximately half of them will be 17 years old and the other half 18. On the other hand, data on new students and those graduating in the same year from medical degrees have been used to establish a snapshot of the situation at any given time, and to analyse whether the Ministry's measure to increase the number of new entry places by 15% could respond to the needs of the SNS. Finally, data on the population of people taking the MIR exam in Spain were used.

Variables and sources of information

To begin with, for the study of enrolment in the university entrance exam, the University Entrance Examination Statistics (EPAU) of the Ministry of Education and Vocational Training and the Ministry of Universities (EDUCAbase) were reviewed. For this purpose, the upper secondary education/vocational training entrance path has been selected. These data have been compared in the analysis with those of the National Institute of Statistics (population by autonomous communities and provinces,

age and sex). This study has been carried out since 2015, in order to provide a wider perspective when it comes to knowing the population and academic evolution of the situation.

With regard to the cut-off marks for the Bachelor's Degree in Medicine, numerous websites have been consulted that provide data on the cut-off marks for each year. In all cases, a detailed study of the data was carried out to verify that there were no significant differences (which in many cases in the temporal analysis depend on when they are published). The range of analysis in this period varies from 2018 to 2022, based on the years in which the SNS has disclosed the results of the allocation of MIR places.

In the study of the admission and graduation and exit of Medicine graduates, the Integrated University Information System (SIIU) has been used as an official source, with data up to the academic year 2020-2021.

Finally, concerning SNS data, the Ministry of Health website (Formación Sanitaria Especializada) has been consulted until the 2021-2022 academic year, which offers a general summary of the MIR entrance exam data.

Procedure and data analysis

Firstly, a descriptive analysis of the extracted data is conducted, describing the data by means of statistical averages and percentages. This analysis provides an overview of the characteristics and distribution of the variables under study. A trend analysis has also been carried out, looking at the data and searching for any upward, downward or seasonal trends in the values over time. Finally, a comparative and integrated study is performed with the different existing situations.

Results

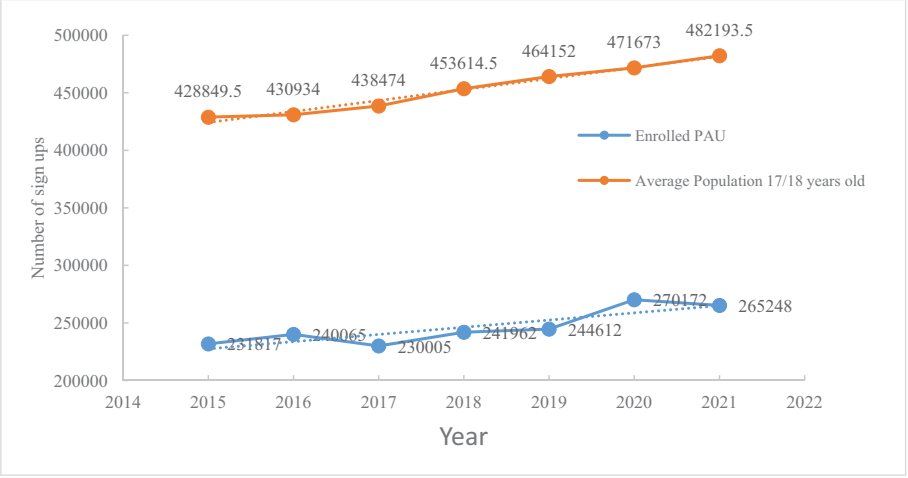
Individuals enrolled in the university entrance examination as a share of the total population of 17-18 year olds

The number of students signing in Spain for the university entrance exam (from Upper Secondary Education/Vocational Training) in the general stage has increased in recent years, with a very significant rise in 2020, coinciding with the first year of the pandemic. Without considering the

Received: 14-11-2022 Accepted: 21-05-2023

exceptional nature of this academic year, in 2021 the increase was also significant with respect to the year before the pandemic, as it increased by 20,636 students, which shows that the number of people who wish to enter university, in absolute quantitative terms, follows a clearly increasing trend year after year (Figure I).

FIGURE I. PAU sign ups and population aged 17 and 18 in Spain (2015-2021)

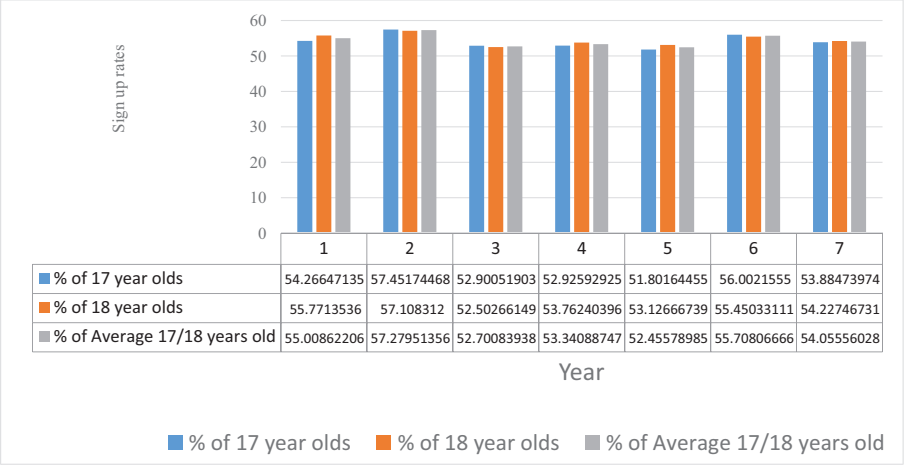


Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). National Institute of Statistics. Population by communities and provinces, age and sex. Prepared by the authors.

In order to analyse this increase, in relation to the change in the total population, absolute figures are compared with the majority population taking these tests (17 and 18 year-olds). Figure II shows the percentages of those signing up over the time analysed with respect to the population aged 17, 18 and the 17 and 18 year-old average, according to what is stated in the methodology section (participants). The data collected from the National Institute of Statistics, show that the percentage of individuals signing barely changes over time and remains constant as a percentage of the general population. Therefore, the share of the average population aged 17 and 18 taking the university entrance exam ranges between 52.5% and 57.3% of the total population (Figure II).

When analysing the data by Autonomous Regions, in absolute terms, we can see in Figure III that most of the students seeking access

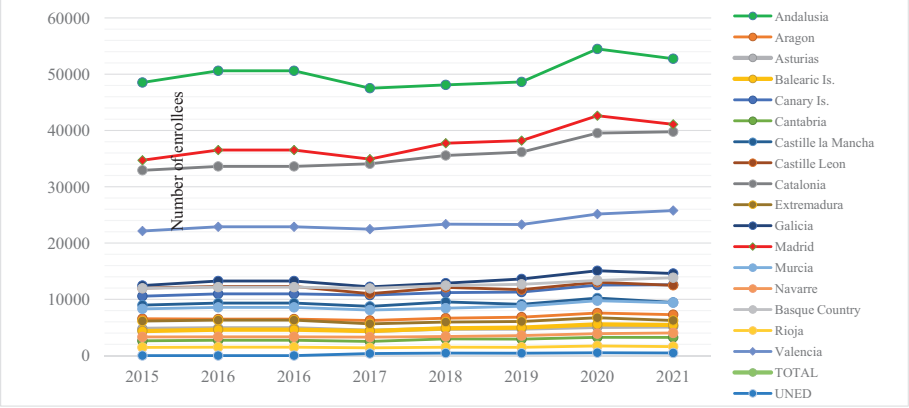
FIGURE II. Percentages of sign ups of total population aged 17, 18 and 17 and 18 year old average population



	1	2	3	4	5	6	7
% of 17 year olds	54.26647135	57.45174468	52.90051903	52.92592925	51.80164455	56.0021555	53.88473974
% of 18 year olds	55.7713536	57.108312	52.50266149	53.76240396	53.12666739	55.45033111	54.22746731
% of Average 17/18 years old	55.00862206	57.27951356	52.70083938	53.34088747	52.45578985	55.70806666	54.05556028

Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). National Institute of Statistics. Population by communities and provinces, age and sex. Prepared by the authors.

FIGURE III. Signups by Autonomous Community



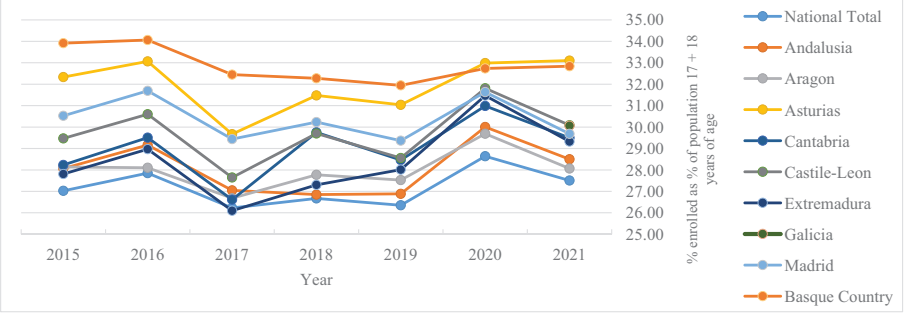
Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). Prepared by the authors.

Received: 14-11-2022 Accepted: 21-05-2023

to the university system come from Andalusia, followed, with a significant difference, by Madrid and Catalonia. The figures in these three are significantly higher than in the rest. These are, admittedly, the most populated Autonomous Communities, but it is clear that most of the students entering the Spanish university system come from these communities.

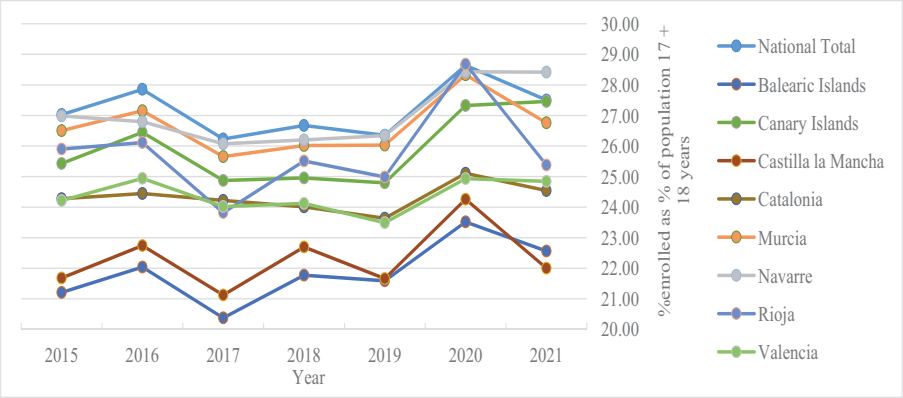
If we consider these data not in absolute terms, but in proportion to the number of sign ups in each Autonomous Community aged between 17 and 18, the data indicate that students aged 17-18 who sign up for the university entrance exam in nine Autonomous Communities (Andalusia, Aragon, Asturias, Cantabria, Castile and Leon, Extremadura, Galicia, Madrid and the Basque Country) are above average, while in eight (Balearic Islands, Canary Islands, Castile and Leon, Extremadura, Galicia, Madrid and the Basque Country) they are above average. (Andalusia, Aragon, Asturias, Cantabria, Castile and Leon, Extremadura, Galicia, Madrid and the Basque Country) are above the average, while in 8 of them (Balearic Islands, Canary Islands, Castile-La Mancha, Catalonia, Murcia, Navarre, La Rioja and Valencia) they are below the average (Figure IV and V).

FIGURE IV. ACs with percentage of sign ups in 17-18 year-olds above the national average



Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). National Institute of Statistics. Prepared by the authors.

FIGURE V. ACs with percentage of signups aged 17-18 below the national average



Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). National Institute of Statistics. Prepared by the authors.

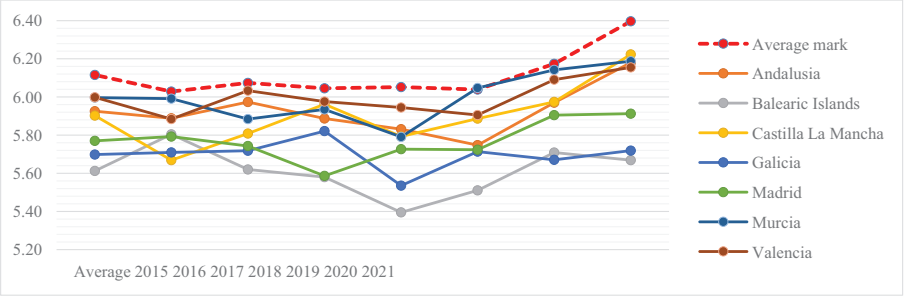
Cut-off marks in university entrance exams

Regarding the evolution of the cut-off marks of the entrance exams (general phase, out of 10 points) in the different Autonomous Communities, it should be noted that since 2015 the average has been rising substantially, standing at 6.12 points in 2021. In the evolution in the years 2015-2021, the Basque Country has the highest average (6.51), while the Balearic Islands has the lowest (5.61), a situation that is maintained in the last year under study.

With regard to the distribution of the scores of the other ACs, it should be noted, as shown in the following figures (VI, VII and VIII), that seven are below the average (Andalusia, Balearic Islands, Castile La Mancha, Galicia, Madrid, Murcia and Valencia); five are above the average (Aragon, Asturias, Cantabria, Basque Country and Navarre); and five are around the national average (Castile Leon, Catalonia, Extremadura, La Rioja and the Canary Islands).

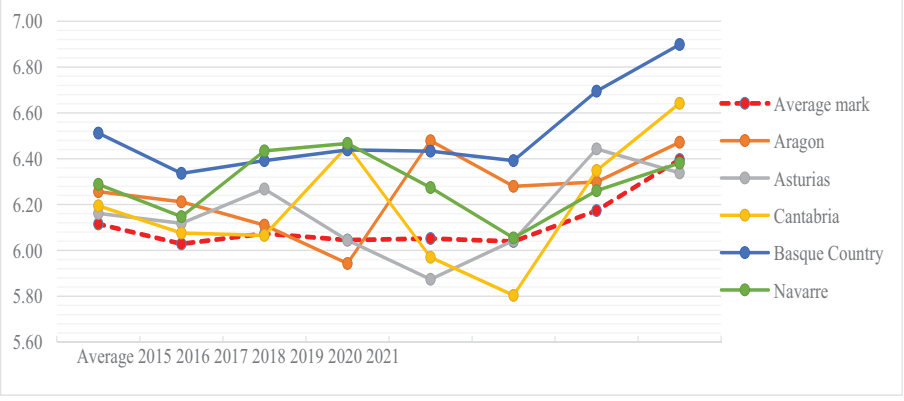
Received: 14-11-2022 Accepted: 21-05-2023

FIGURE VI. Average general stage entrance examination marks by Autonomous Community (below the national average)



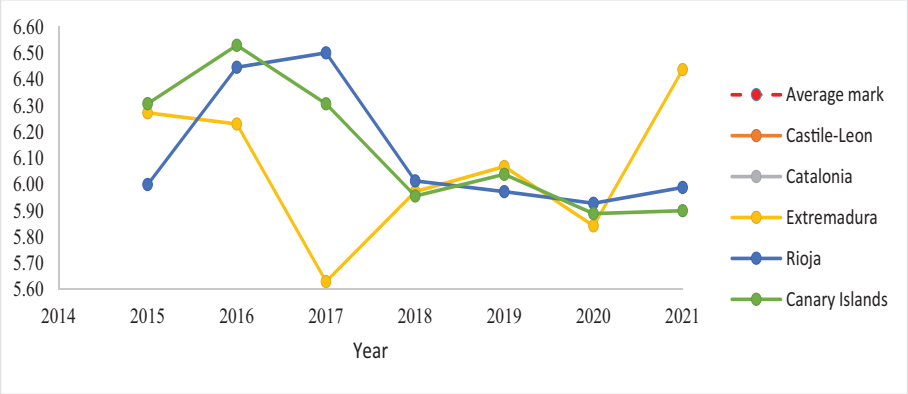
Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). Compiled by the authors.

FIGURE VII. Average general stage entrance examination marks by Autonomous Community (above the national average)



Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). Prepared by the authors.

FIGURE VIII. Average general stage entrance examination marks by Autonomous Community around the national average



Source: Ministry of Education and Vocational Training and Ministry of Universities. EDUCAbase. University Entrance Examination Statistics (EPAU). Compiled by the authors.

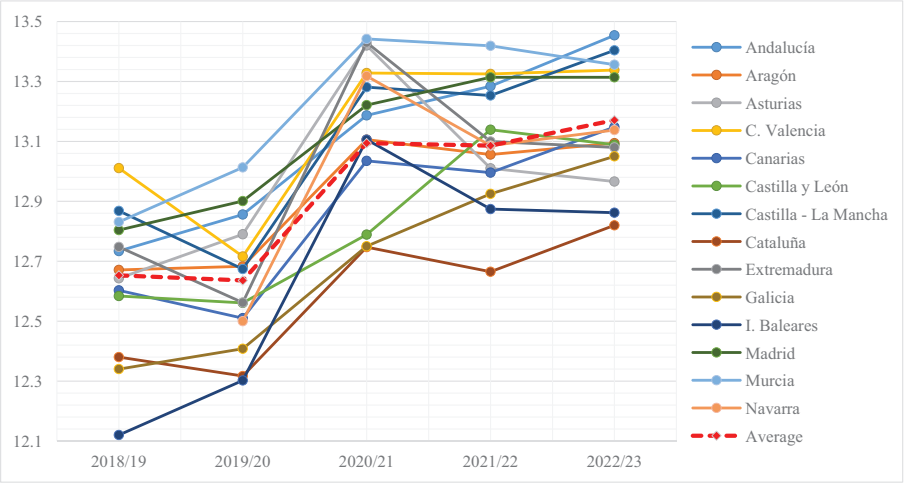
Cut-off marks for the Bachelor's Degree in Medicine

In general terms, the cut-off marks for Health Sciences degrees are the highest in the system. Since the 2018-2019 academic year, the average mark in public universities for entry to the Bachelor's Degree in Medicine has risen from 12.653 to 13.171 (out of 14) in the 2022-2023 academic year. This gradual rise, in the period under study, shows a significant jump in the academic year 2020-2021, the year of the pandemic, which also coincides with the increase in the number of students signing up for the entrance exam, where the average mark increased by half a point (Figure IX).

This analysis does not include private universities, for which in it is not possible to obtain this information many cases, as it is usually not a determining factor for admission to Medicine studies at these institutions. However, if we look at the regions, it is worth noting that there are some Autonomous Regions that are consistently above the average. This is clearly the case in Andalusia, which seems to be a growing preference for those wishing to study Medicine.

Received: 14-11-2022 Accepted: 21-05-2023

FIGURE IX. Average entry marks in the Medicine Degree by Autonomous Region



Source: Educaweb.com; Notasdecorte.es; Distrito unico andaluz (notas de corte) et al.

In this sense, as Figure IX shows, although the general trend is a high average mark for access to Medicine studies, as it happens in more than 50% of the universities studied throughout the period, we can see that it is in the academic year 2020-2021 were we can find the highest number of universities exceeding the national average, with 9 of the 14 universities (64.3%). It is not until the academic year 2021-2022 that the number drops to 5, i.e. 35.7% of the universities included in the study.

The highest entry grades in Spain, which exceed the cut-off mark of 13.3, are found in the Autonomous Communities of Madrid (13.314), Valencia (13.338), Murcia (13.356), Castile-La Mancha (13.404) and Andalusia (13.454), the latter being the highest. We can see that in the case of the Andalusian universities, the progression towards a higher cut-off mark is always increasing between the academic years 2018-2019 and 2022-2023, going from being the eighth highest in the academic year 2020-2021 to, in two academic years, becoming the most demanding autonomous community at the time of admission to the studies of Medicine.

On the other hand, although the academic year 2019-2020 is marked by a decrease in the entry mark in 50% of the autonomous communities analysed in this study, the academic year 2020-2021 is a turning point in the level of demand for the new entry of students to the degree of Medicine in our country. Thus, the entry grade continues to rise in the regions of Andalusia, Castile-La Mancha and, very slightly, Valencia, despite being the ones with the highest cut-off marks, and in Galicia. The rest of the Autonomous Communities recorded a decrease in the entrance examination marks. It is the 2021-2022 academic year that will see an increase in this cut-off mark in 50% of the autonomous communities that incorporate Medicine students to the system.

Catalonian universities are the ones with the lowest cut-off marks for admission to Medicine studies throughout the historical sequence analysed, with the highest cut-off mark of 12.82 recorded in the academic year 2022-2023, which is 0.36 points below the national average and 0.63 points below the highest in the system in that academic year.

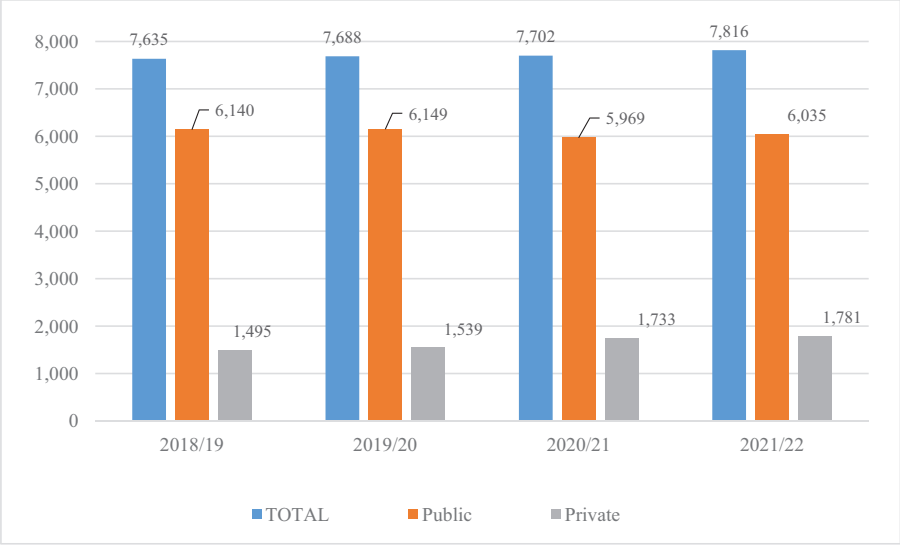
New students/graduates in the Bachelor's Degree in Medicine

The number of students enrolled as new students in the Bachelor's Degrees in Medicine throughout Spain has been increasing in recent years, although such increase has not taken place in the same extent in public and in private universities. Figure X, shows that private universities have been progressively increasing the number of newly enrolled students, whereas in public universities the number of places has decreased by 105 from 2018 to 2021.

By Autonomous Communities, Madrid, Catalonia, Andalusia and Galicia have the largest number of newly enrolled medical students, and the other Autonomous Communities lie far behind. Additionally, in these three cases, the number of newly enrolled students has decreased considerably in the case of Catalonia, in Andalusia where it is below the 2018 academic year, as well as in Galicia, where it is also slightly below the 2018 figure (Figure XI).

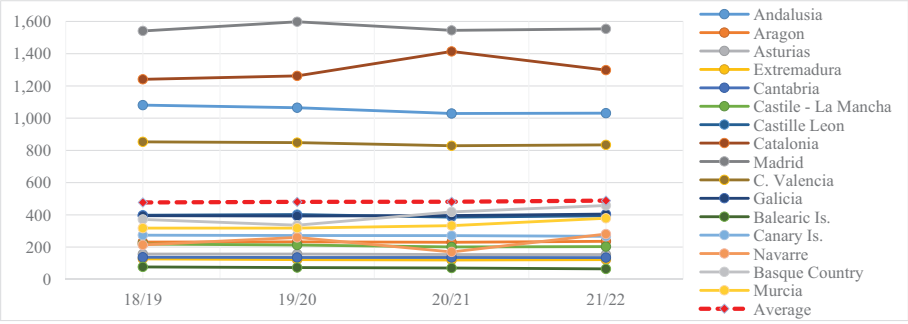
The rest of the universities have maintained the enrolment of new entrants practically constant since the 2018-2019 academic year, with two exceptions: The Basque Country, where enrolment has increased exponentially since the 2019-2020 academic year, and Navarre, where enrolment fell in the 2020-2021 academic year, even though the Public

FIGURE X. Evolution of new enrolments in the Degree in Medicine



Source: Integrated University Information System (SIIU). Compiled by the author.

FIGURE XI. Evolution of new enrolment in the Degree in Medicine by Autonomous Community

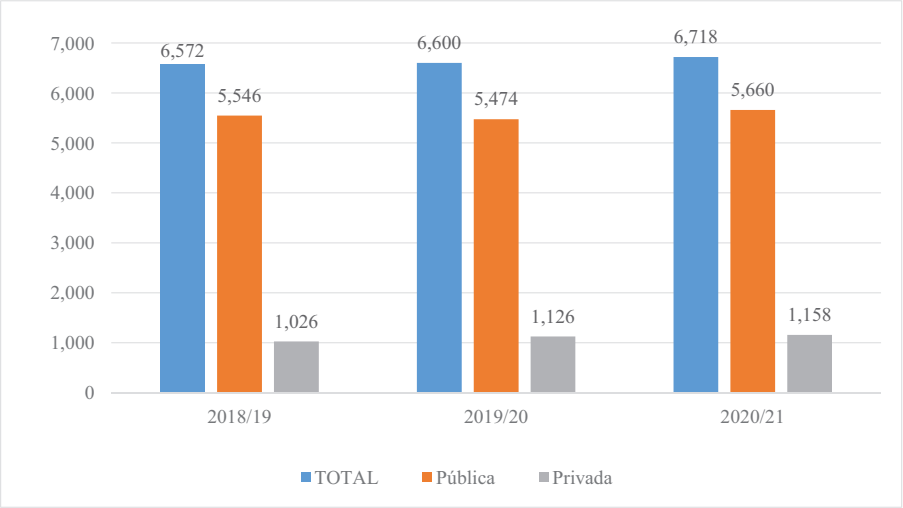


Source: Integrated University Information System (SIIU). Compiled by the author.

University of Navarre had introduced the Bachelor's Degree in Medicine for the first time in 2019.

With regard to the number of medicine students graduating annually from the Spanish university system (Figure XII), up to the academic year

FIGURE XII. Evolution of graduates of the Degree in Medicine



Source: Integrated University Information System (SIIU). Compiled by the author.

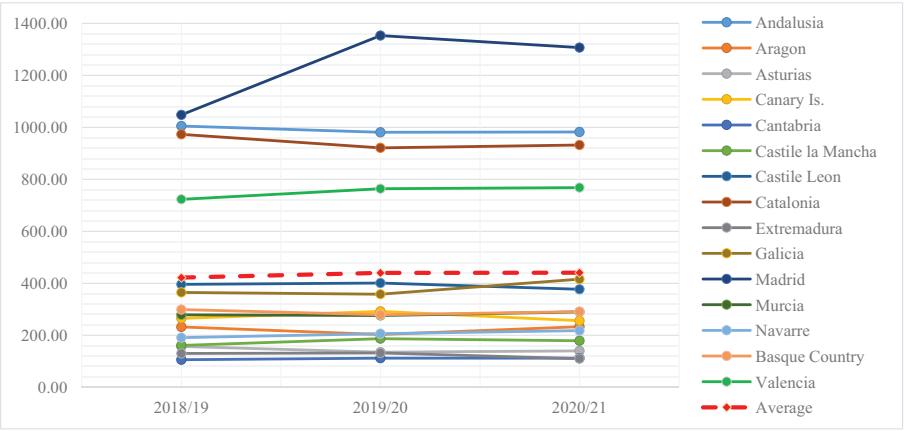
for which the data are available, this has remained constant (6,572, 6,600 and 6,718), with around one thousand fewer students graduating than the number entering university every year. If the average number of new entrants over these years is 7,710, the average number of graduates over these years is 6,630. But the proportion in this case is different if they have done so in public or private universities. In the case of public universities, an average of 91% of those entering each year graduate. In the case of private universities, the average is 67%.

In the analysis by Autonomous Regions, as in the case of new entrants, four are the Autonomous Regions with the highest number of graduates in the years under study. In relation to this aspect, of the four, Andalusia is the one with the smallest difference between the number of new entrants and graduates (76, 84 and 47 students per academic year), whereas in Catalonia this difference is larger and keeps increasing, with the difference in the number of students being 268, 342, and 468, respectively in the 2018-2019, 2019-2020 and 2020-2021 academic years, probably due to the increase in the number of new entrants. In

Received: 14-11-2022 Accepted: 21-05-2023

the other range, as in the case of new enrolments, the number of graduates remains constant, with the exception of Castile Leon, the Canary Islands and Extremadura, where there was a slight decrease in the last year. In these cases, the difference between the number of new entrants and graduates in these years is practically nil, and in some years it is even higher (Figure XIII).

FIGURE XIII. Evolution of graduates of the Degree in Medicine by Autonomous Community



Source: Integrated University Information System (SIIU). Compiled by the author.

Summary of MIR exam data

According to the general summaries of the MIR calls published in the website of the Ministry of Health (Specialised Healthcare Training) over the last four years, as shown in the following table, the number of places offered by the health system has progressively increased by almost 1,400 places since 2018. While this increase has taken place, the same has not happened with the number of applications submitted and admitted to the exam. Since 2020-2021, there has been a clear decrease in the number of individuals taking the MIR entrance exam.

TABLE I. Summary of MIR exams

ADMISSION STAGE	2018/19	2019/2020	2020/2021	2021/2022
Applications submitted	16,582	16,964	15,166	13,080
Number of candidates admitted to the examination	15,700	16,263	14,452	11,827
SELECTION STAGE	**2018/19**	**2019/2020**	**2020/2021**	**2021/2022**
No, of candidates taking the exam	14,187	14,986	13,332	11,827
Candidates eliminated after cut-off mark	2,983	2,796	2,527	1,895
Candidates with order number	11,204	12,174	10,805	9,932
AWARD STAGE	**2018/19**	**2019/2020**	**2020/2021**	**2021/2022**
Places offered	6,797	7,615	7,988	8,188
Places allocated	6,796	7,615	7,987	8,095
Unallocated places	1	0	1	93

Source: Ministry of Health. Spanish Government. Compiled by the author.

Of those who were finally admitted (approximately 90%) and took the exam, the share has been increasing, standing at around 83% at in the last year under study, In addition to this fact, the gap between the number of applicants with a qualifying number and the number of places offered has been narrowing, being the smallest in the 2021-2022 call for applications, where the difference was 1,744, compared to the 2019-2020 call for applications, where the difference between those who passed the cut-off mark and the places offered was 4,559,

Finally, it should be noted that in two of the years under study, and according to the published data, one vacancy was not awarded respectively, as compared to the 93 vacancies that have not been awarded in 2021-2022. All of them belong to the speciality of Family and Community Medicine.

Conclusions

As it has been exposed, the road leading to the medical profession is a complex one, which entails access to university. And it is significant that

Received: 14-11-2022 Accepted: 21-05-2023

in order to respond to the deficit of professionals in the SNS, the university has been called upon to increase the number of places in medicine by up to 15%. This decision has led us to consider a general reflection on access to the system, focusing on the case of the Bachelor's Degree in Medicine. We are aware that the study has been approached with too much ambition, and that it therefore poses serious limitations.

In recent years, the number of people taking the university entrance exam has been increasing steadily, but, as we have seen, this is a consequence of the fact that the population has also been growing proportionally. The exception to this situation happened in 2020, the year in which the pandemic was declared, when the trend was broken, with a significant increase in the number of people wanting to enter university. When the analysis is made in proportion to the relevant population in each Autonomous Community in that age range, it is striking that the data show a situation in which, over time, in some Autonomous Communities the rate of people applying in proportion to those who can access university is above the average, while in others it is constantly below. What aspects could be defining these differences between Autonomous Communities? This situation would undoubtedly merit an analysis, which goes beyond the scope of this paper, linking this situation to other variables (e.g. immigration, urban vs. rural population, socio-economic and language factors, etc.).

Two questions could be raised in this regard. The first of these refers to the unquestionable interest of the study of the effects of the pandemic on access to university. There are many questions that can be raised in this regard: did the change in the examination model cause this increase, was it a generalised desire or was it focused on the desire to access the health professions, which played such a crucial role throughout this period? The second question links the increase in the number of people entering university with the cut-off marks, which have also increased progressively over the years. But it cannot be ignored that the cut-off marks are the result of the number of places that universities offer, and of the social demand. Therefore, if universities have not increased their offer of places in the years under study, this may have led to a general increase in the cut-off marks.

On the other hand, the results of the 17 existing entrance exams in Spain show a heterogeneity in the scores that has a negative effect on equal opportunities for access in general, and specifically for access to

medical studies. If the university entrance exam is understood as a single gateway to the system, it makes no sense that, over a sustained period of time, some regions are below the average, while others are above the average. Instead, they should all be around the average. Another aspect that deserves further study is the fact that the highest cut-off mark for studying medicine, and with an upward trend, is found in Andalusia, while the lowest one Catalonia. The language issue could be considered among the causes of the latter situation, but this would also merit further study.

In the analysis of the heterogeneity of grades, it would be interesting to introduce a study of the university entrance examination itself, since the literature has shown that there are numerous aspects that require further standardisation. One example of how to correct differences could be to review the grading scale used (Veas et al., 2020). It would be very relevant, and it is a major limitation of this work, to carry out a study of successful access systems in other countries, and extrapolate those issues that could help to improve the system.

On the other hand, the progressive increase in the cut-off marks for Medicine studies shows that the trend will continue to rise in the coming years, in some cases, such as in Andalusia, it will approach 14 points. This aspect could be corrected with the 15% increase in the number of places proposed by the Ministry for the coming academic year, which will be confirmed in the near future. But while this may be a modulating measure for the cut-off mark, it is not yet clear that it could be a corrective measure for the problem of the lack of medical professionals, since the data show that the correlation between the two remains constant, with an average of 91% of graduates in the case of public universities compared to those enrolled that year, and an average of 67% in the case of private universities. In other words, it is possible that these figures are linked to the drop-out rate, an aspect that should be studied, or to the creation of new universities. Notwithstanding, it is another major limitation of this study not to have carried out a longitudinal study in this case.

Finally, with regard to the MIR entrance exam, it might seem obvious that the number of places on offer meets the numbers of the university system. In 2018-2019, 2019-2020 and 2020-2021 respectively 6,572, 6,600 and 6,718 medical students have graduated from university and the number of places for access to the MIR in these years has been: 6,797, 7,615

Received: 14-11-2022 Accepted: 21-05-2023

and 7,988. However, if we analyse the number of physicians taking the exam, which is practically twice as high, we can understand that there is a population of graduates (and foreigners) who have not entered the SNS yet, and continue to take the exam in successive call. Additionally, it is striking that in the 2021-2022 call for applications, 8,188 vacancies have been offered, the highest number in recent years. And for the first time, according to the Ministry's official data, in the annual summary published, 93 places appear "pending awarding" (in previous years, 2). All of them in the speciality of "Family and Community Medicine", Primary care. The data presented here on the number of graduates and the number of applicants do not initially seem to fit this situation. This could be due to the fact that it is not a speciality in demand and could also be a response to the high number of people who take the MIR exams each year compared to the number of places offered. Therefore, it is possible that there are people already in the health system in the speciality of "Family and Community Medicine" who try to obtain a place in another speciality every year and leave primary care. It is therefore worth considering whether there is a shortage of physicians in general or a shortage of specialists, and rethink the need to increase the 15% of new places in medicine in the university system, as this does not seem to be the solution to the problem.

In this scenario, the increase in the number of places in medical programmes could increase the number of graduates, but this needs to be proven. Increasing the number of graduates would increase the number of MIR places in all specialties, but what about Family and Community Medicine? The struggle in the system regarding access to medical studies does not seem to satisfy the professional outlets. Perhaps in the area of training future doctors, the university could contribute something more than just merely increasing the number of new places in this degree. Access to specialisation for medical graduates could be rethought, separating training in primary care, either by creating a specialised degree in this area, or by creating a professional-leading postgraduate programme for nursing graduates. This would make it possible to organise the system, admit to the field of medicine those people who could not gain access because of the cut-off mark, and regulate the tension that apparently arises among current graduates to fill the vacancies on offer. In this case, once again, the university, as a public service, could provide the solution.

Bibliographic References

Baillès E., Girvent M., Moyano E., & Pérez, J. (2020). Relationship between academic record and result in the MIR test in medical studies at UPF-UAB. *FEM: Revista de la Fundación Médica*, Nº *23*(4)23, vol. 4, 205-209. https://dx.doi.org/10.33588/fem.234.1069

Baladrón Romero, J., Sánchez Lasheras, F., Peñalver San Cristóbal, C., García Guerrero, A., Romero Ladreroe, J.M., Villacampa, T., Curbelo, J., & Jiménez Fonseca, P. (2022). Results obtained in the MIR test according to academic scale. Convocatorias of 2019 and 2020. *Investigación en Educación Médica,* Nº 11, vol.(43), 51-62. https://doi.org/10.22201/fm.20075057e.2022.43.22420

Boal, N., Bueno, C., Lerís, M. D., & Sein-Echaluce, M. L. (2008). Mathematical skills assessed in university entrance exams. A study in several Spanish public universities. *Revista de Investigación Educativa*, Nº 26, vol.1, 11-23.

Bowen, G. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal, 9*(2), 27-40. DOI: https://doi.org/10.3316/QRJ0902027

Cut-off marks for Medicine (30 October 2022). https://sites.google.com/site/notasdecorte/notas-de-corte-de-medicina-2020-21

Cut-off marks for Medicine (30 October 2022). https://sites.google.com/site/notasdecorte/notas-de-corte-de-medicina-2019-20

Del Sol, J.R. (1974). La selección, problema candente en la enseñanza universitaria. *Revista de Educación* nº 230. *Monográfico Acceso a la Educación Superior*, 5-20.

Echevarría, A., P., (10 October 2022). We do not need more students, but more MIR places and better conditions. *Revista médica*. https://www.redaccionmedica.com/virico/noticias/-no-necesitamos-mas-estudiantes-sino-mas-plazas-mir-y-mejores-condiciones--9974

Educaweb (30 October 2022). Notas de corte de corte de Grado de Medicina. https://www.educaweb.com/notas-corte/grados/ciencias-salud/grado-medicina/; https://notasdecorte.es/medicina

Faura-Martínez, U., Lafuente-Lechuga, M., & Cifuentes-Faura, J. (2022). Territorial inequality in Selectividad? Analizando la asignatura de matemáticas en Ciencias Sociales. *Revista de Investigación Educativa,* Nº 40, vol.(1), 69-87. DOI: http://dx.doi.org/10.6018/rie.42484

Received: 14-11-2022 Accepted: 21-05-2023

Integrated University Information System (SIIU) (30 October 2022). *Students enrolled and graduates.* https://www.educacion.gob.es/siiu

Jiménez García, E., Arroyo Resino, D., Hurtado-Martín, M., Ruiz-Lázaro, J., Sánchez-Munilla, M., Illana Vicaria, J.J., & González Barbera, C. (2021). The university entrance mark as a predictor of performance in the first year of the degree course: teaching degrees versus other careers. *Revista de Educación*, nº 393, July-September, 29-154.

Junta de Andalucía (30 October 2022). Distrito Único Andaluz. Cut-off notes. https://www.juntadeandalucia.es/economiaconocimientoempresasyuniversidad/sguit/?q=grados&d=g_not_cor_anteriores_top.php

Laurence, C.O., Zajac, I.T., Lorimer, M., Turnbull, D.A., & Sumner, K.E. (2007) The impact of preparatory activities on medical school selection outcomes: a cross-sectional survey of applicants to the University of Adelaide Medical School in 2007. *BMC Medical Education*, No. 13, DOI: 10.1186/1472-6920-13-159.

Law 44/2003, of 21 November [Official State Gazette, 280] on the organisation of the health professions of *Of* 22 November 2003.

Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario [Boletín Oficial del Estado, 70] núm. 70, de 23 de marzo de 2023.

Organic Law 3/2020, of 29 December, [Official State Gazette, 340] amending Organic Law 2/2006, of 3 May, on the Education of. *Of* 30 December 2020.

Mengual, E., Albarracín, L., Muñoz-Escolano, J.M., Oller-Marcén, A.M., & Gorgorió, N. (2019). Design of criteria to reduce variability in the grading of Mathematics exams in university entrance exams. *PNA Revista de investigación en didáctica de la matemática*, Nº13, vol. 2, 62-83.

Ministry of Education and Vocational Training and Ministry of Universities (30 October 2022). *EDUCAbase. Statistics of the University Entrance Examination (EPAU).* http://estadisticas.mecd.gob.es/EducaJaxiPx/Tabla.htm?path=/Universitaria/PAU/PAU21//l0/&file=PAU0105.px&type=pcaxis&L=0

Ministry of Education and Vocational Training (30 October 2022). Registry of Universities, Centres and Degrees (RUCT). https://www.educacionyfp.gob.es/servicios-al-ciudadano/catalogo/centros-docentes/servicios-generales/ruct.html

Ministry of Health. Formación Sanitaria Especializada (30 October 2022). *General summary of the call for entrance exams (*2021, 2020, 2019

and 2018). https://fse.mscbs.gob.es/fseweb/view/public/datosanteriores/resumenGeneral/busquedaConvocatoria.xhtml

Ministry of Universities (2022). *Datos y Cifras del Sistema Universitario Español (Publication 2021-2022)*, Ministry of Universities. https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/DyC_2021_22.pdf

National Institute of Statistics. *Population by communities and provinces, age and sex*. (30 October 2022). https://www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p04/provi/l0/&file=0CC.AA.003.px&L=0

Muñoz-Repiso Izaguirre, M., & Murillo Torrecilla, F.J. (1997). Los resultados de la selectividad actual: algunas cuestiones a debate. *Revista de Educación, Monográfico Las pruebas de acceso a la universidad*, Nº 314, 29-48.

Murillo Torrecilla, F.J. (1997). Análisis de las pruebas que conforman la selectividad. *Revista de Educación, Monográfico Las pruebas de acceso a la universidad*, Nº 314, 49-62.

Patterson F, Roberts C, Hanson M.D. et al. (2018). Ottawa consensus statement: selection and recruitment to the healthcare professions. *Medical Teacher.* No. 40, vol. 11, 1091-1101.

Peña Vera, T. (2022). Etapas del análisis de la información documental. *Revista Interamericana De Bibliotecología, 45*(3), e340545. https://doi.org/10.17533/udea.rib.v45n3e340545

Pérez-Cárceles, M.C, & Martínez-Martínez, M. (2019): Dimensions of academic performance in the University Entrance Examination in Spain. *Revista de Estudios Regionales*, Nº 116, 67-91.

Royal Decree 640/2021, of 27 July, [Official State Gazette, 179] on the creation, recognition and authorisation of universities and university centres, and institutional accreditation of university centres. Of 28/07/2021.

Royal Decree 822/2021 of 28 September [Official State Gazette, 233] establishing the organisation of university education and the procedure for quality assurance. 29 September 2021.

Ruiz de Gauna Gorostiza, J., & Sarasua Fernández, J. (2011). Do results improve with the new Selectividad system? *Bordón: Revista de pedagogía*, vol. 63, Nº 4, 111-122.

Ruiz-Lázaro, J. (2021). *University entrance in Spain. Comparative analysis of the common tests by autonomous communities* (Doctoral thesis). Complutense University of Madrid. Del Sol, J.R. (1974). La selección,

problema candente en la enseñanza universitaria. *Revista de Educación* nº 230. *Monográfico Acceso a la Educación Superior*, 5-20.

Ruiz-Lázaro, J., & González Barbera, C. (2017). Análisis de la prueba de Lengua Castellana y Literatura que da acceso a la universidad: comparación entre las comunidades autónomas, *Bordón: Revista de pedagogía*, Nº 69, vol. 3,175-195,

Ruiz-Lázaro, J., González Barbera, C., & Gaviria Soto, J.L. (2021). English tests for university entrance. A comparison between Autonomous Communities. *Educación XXI*, Nº 24, vol.1, 233-270.

Sentí M., Pérez J. and Baños J.E. (2016). Predictors of results in the MIR test in public universities. Analysis of the 2008-2014 cohort. *FEM: Revista de la Fundación Médica,* Nº19, vol. *319*(3), 155-60. https://dx.doi.org/10.33588/fem.193.842

Veas, A., Benítez, I., Navas, L., & Gilar-Corbí, R. (2020). Comparative analysis of university entrance exams under the construct comparability approach. *Revista de Educación*, Nº 388. April-June, 65-84.

Vicenç Martínez Ibáñez, V., Carbajo Arias, P., Pérez Fernández-Turégano, C., & Verónica Ledo Cosqui, V. (2022). Reflections on the MIR (2020-2021). *Revista Española de Salud Pública*, Nº 1- 6. https://www.sanidad.gob.es/biblioPublic/publicaciones/recursos_propios/resp/revista_cdrom/Suplementos/Perspectivas/perspectivas23_martinez_carbajo_perez_ledo.pdf

Contact address: Cristina Yanes Cabrera. Universidad de Sevilla, Facultad de Ciencias de la Educación, Departamento de Teoría e Historia de la Educación y Pedagogía Social. Despacho 5.111 (Bloque II, 5ª planta). c/ Pirotecnia s/n. C.P.: 41013, Sevilla (Spain). E-mail:yanes@us.es

Received: 14-11-2022 Accepted: 21-05-2023

Reseñas

Lluch Molins, L. and Cabrera Lanzo, N. (Coords.) (2022). *Competence of learning to learn and self-regulation at the university*. Barcelona: Octaedro. 130 pp. ISBN: 978-84-19312-60-0

This book describes five practical applications of peer assessment among students in subjects from different university degrees: Teaching Degree, Master's Degree in Education and ICT, Archeology Degree, and Biotechnology Degree. It is aimed, therefore, at teachers working in higher education.

The first chapter is an updated review of the concept and impact of assessment among students in formal learning contexts. Laia Lluch and Nati Cabrera provide the theoretical foundations that support the formative power of peer feedback. The researchers explain how this tool favors the self-regulation of the student body.

Below are the five experiences described by the teachers who designed and implemented them in their subjects. All the descriptions follow the same outline of contents. First, the subject is located in its context and the activities to which the peer feedback is applied are defined. The methodology is justified below. Transversal skills developed that affect self-regulation are indicated. Afterward, the tasks of the students and the teacher are detailed. It describes how these tasks were organized and the feedback among students. Finally, the advantages and disadvantages found are described. This way of presenting the work allows the readers to understand how to apply peer feedback in other contexts.

The first experience is told by Carles Lindín. He applied peer feedback in the Digital Literacy subject of the Teaching Degree at the University of Barcelona (UB). The activity was individual and consisted of making an e-portfolio on the incorporation of technology in education. Two feedback loops were implemented between pairs of students. This allowed two enhancement opportunities for each student. The students gave feedback based on the task evaluation criteria established by the teacher.

The second experience was given in an optional subject of the Archeology Degree at the UB taught by Marta Sancho. The activity was carried out in groups of three to five students. It consisted of making a mental map. The feedback was provided by each group of students to another group that, after receiving it, could implement the suggestions of their peers. She used the MIRO tool. It allowed students to make the map, communicate through the chat, and register the feedback.

The third experience was implemented at the Universitat Oberta de Catalonia's Master's in Education and ICT in the online modality. It is described by Nati Cabrera, Lourdes Guàrdia, Maite Fernández and Marcelo Maina. The task was individual and consisted of writing an essay. A chatbot, a doubts forum, and another forum containing peer feedback were used. Each student provided feedback to a peer, which was randomly assigned, before final delivery.

The fourth experience is told by Laia Lluch. It was given in a subject of Teaching degree at the UB. The activity was in groups and gamified. The students had to design a complex, gamified, and inclusive task from the perspective of the integrated curriculum. The feedbacks were individual and two loops were implemented. To carry them out, Lluch established evaluation criteria both for the design of the task and for the quality of the feedback.

The last experience was given in the Genetic Analysis subject of the second year of the Degree in Biotechnology at the UB. It was taught by David Bueno. The group consisted of 80 students. This subject is taught simultaneously in other university degrees. The activities and evaluation systems are rigidly established by the faculty, and all tasks are individual. Bueno, however, managed to introduce peer feedback and cooperative learning in a context where it seemed impossible: taking a personal multiple-choice exam. That is, he trained his group of students in self-regulation competence in conditions where this is not expected.

The last chapter collects the conclusions of the coordinators of this work. It is explained that digital tools improve any didactic methodology because they facilitate interaction and the recording of evidence. They argue that peer feedback makes the tasks more motivating and that the students take ownership of the assessment criteria.

Salvador Javier Ros Turégano

Norton, J. and Buchanan, H. (Eds.) (2022). *The Routledge Handbook of Materials Development for Language Teaching*. Routledge: New York. 560 pp. ISBN: 9780815382577

It may be questioned why there is a need for an academic book on this subject, given that materials development is typically focused on practical application and there are already many materials available to professional writers. While these existing materials can serve as a useful reference point, they may also encourage the creation of new materials that simply mimic what is already available in response to market demands. However, contrary to some opinions, materials development is not solely a practical pursuit but rather a field of study that involves principles and procedures for its development. Based on this definition, Julie Norton and Heather Buchanan collaborated to produce the current handbook with the goal of combining theory with practice by gathering insights from more than forty contributors from various international contexts.

The Routledge Handbook of Materials Development for Language Teaching, edited by Julie Norton and Heather Buchanan, is a comprehensive guide to the development of materials for language teaching. It aims to combine theory with practice by gathering insights from over forty contributors from around the world. This book covers a wide range of topics and challenges the idea that materials development is purely practical by emphasizing the connection between theory and practice. In comparison to other recent books on the topic, this handbook offers broader coverage and is not limited to the TESOL field.

Despite its title, this handbook is not just a manual but rather a comprehensive guide with thirty-four chapters organized into nine parts. It discusses four main ideas, each covering different aspects of materials development. The first idea provides an overview of the field of materials development and its history. The second idea focuses on the theoretical underpinnings of materials development, including principles of language learning and teaching. The third idea covers practical considerations in materials development, such as needs analysis and evaluation. Finally, the fourth idea explores emerging trends in materials development, such as digital technologies and open educational resources. Each chapter is written by a different expert in the field and provides valuable insights into their particular area of expertise.

The first part of this handbook is comprised of three chapters that outline the changes and advancements in language teaching materials. Chapter 1 details the rapid progress of materials development. Chapter 2 introduces criteria for evaluating technology-mediated materials. In contrast, Chapter 3 reviews the current status between theory and practice with regard to the four skills, vocabulary, and grammar. This chapter aligns more closely with the editors' objective as it examines critical issues, outlines theoretical principles, justifies rationales persuasively, and highlights challenges involved in producing theoretically informed materials.

The second part of the handbook discusses contemporary debates in materials development in six different chapters. These chapters tackle contentious topics like the significance of international course books (Chapter 4), the advantages of authentic materials (Chapter 5), language collection for materials writing (Chapter 6), modern approaches to ELT materials (Chapter 7), cultural content of materials (Chapter 8), and representation in course books (Chapter 9). These discussions warrant additional study, as the field of materials development evolves. However, there are still some outstanding issues that demand attention but were regrettably not included in this publication. For instance, humanizing materials are required, and undue reliance on computerized language-learning tools needs to be avoided.

In part three of this book, there are three chapters that examine the correlation between research and the development of materials. Chapter 10 focuses on studying the development of materials' content. The use of diverse research approaches to guide materials development is covered in detail in Chapter 11. Finally, Chapter 12 examines the advantages of utilizing spoken corpora for the creation of educational resources and material development. However, it is recommended that this handbook also takes into account stakeholder cooperation to carry out a longitudinal investigation on the results of using various materials based on SLA principles.

The materials for language learning and skill development are also covered in five different chapters in part four of this handbook. Grammar instruction with an emphasis on content and scope is the main topic of Chapter 13. Effective strategies for teaching vocabulary are covered in Chapter 14. In order to build proficiency in speaking, Chapter 15 examines various linguistic features for developing speaking skills. Chapter 16

emphasizes developing a realistic reader-text interaction for the improvement of reading skills, while Chapter 17 focuses on product, process, and sociocultural activity in developing writing skills. However, it is noted that a chapter on materials for developing listening skills is missing from this section.

In essence, part five of this handbook provides guidelines and procedures for instructors and material designers to assess and modify instructional materials. It is crucial for educators to evaluate the effects of the materials they employ on their students and make the necessary adjustments. Chapter 18 covers the two phases of competence needed for examining and evaluating teaching materials, while Chapter 19 explores approaches to adapting materials for future direction in global English. I suggest that the handbook should encourage a more inclusive approach that involves learners adapting their own materials with or without their teacher's guidance.

There are seven chapters in part six exploring the materials development in various contexts. For example, chapter 20 discusses the primary course book series in Turkish, Chapter 21 explores versioning course books for broader audiences, Chapter 22 reveals the writing skills required for English for Academic Purposes (EAP), Chapter 23 makes the case for more ESOL-specific materials that take learners, context, and pedagogical factors into account, Chapter 24 addresses the need for English-medium materials in Europe and Sub-Saharan Africa, Chapter 25 concentrates on writing course books for monolinguals and teenagers in Spanish, and Chapter 26 considers the difficulties in writing materials for an English-speaking environment.

The most noteworthy section of this handbook, in my opinion, is part seven which includes two chapters about incorporating technology into language-learning materials. Despite the fast-paced and fleeting nature of technology, these chapters provide valuable information. Chapter 27 examines the development of blended learning materials and the ethical standards for producing both commercial and local content. Chapter 28 examines the current trend of mobile devices and how writers can evaluate the effectiveness of using mobile materials. Technology has provided teachers and learners with more options for teaching and learning outside of the classroom, especially during the pandemic when many students have access to their mobile phones but not their physical classrooms.

Part eight of this handbook is also commendable for including three chapters that give voice to underrepresented perspectives in materials development. These skilled industry professionals provide useful insights for individuals looking for a deeper comprehension of materials development. For example, Chapter 29 emphasizes the part that students play in developing and modifying commercial and other sorts of materials. Chapter 30 examines a recent development in publishing where authors and editors work together to create materials for for-profit publications. Finally, Chapter 31 offers suggestions for editorial training as well as potential directions for ELT editors and publishers in creating instructional resources for the publishing sector.

The handbook includes three chapters in Part Nine that focus on professional development and materials writing. Chapter 32 shares insights into print materials, digital media, and real-world writing projects. The ideas in Chapter 33 indicate ways for teachers to use the content from course books for professional development and how publishers could contribute. Finally, Chapter 34 suggests a creative curriculum to train authors in the creation of materials through multiple activities.

In conclusion, some recommendations based on the handbook that could improve future editions should be made. Firstly, the issue of humanizing materials should be addressed as it is crucial for affective and cognitive engagement in the age of artificial intelligence. Secondly, new technologies should not replace meaningful face-to-face classroom interaction and reliable resources should be ensured. Tomlinson (2020) emphasizes that materials development research should focus on making discoveries that benefit both teachers and learners, although concrete empirical evidence may be challenging to present. Thirdly, listening skills should be given the same weight as reading skills when developing materials for skills development.

Despite this, I believe that the standout features of this book lie in its comprehensive coverage and practical guidance. The book covers a wide range of topics related to materials development for language teaching, including theoretical frameworks, practical considerations, and case studies. Furthermore, it provides practical guidance on how to develop effective materials for language teaching, including tips on designing tasks, selecting texts, and incorporating technology. This collection is grounded in research and draws on current theories and best practices in materials development for language teaching. It includes useful resources such as

checklists, templates, and sample tasks that can be adapted for different contexts and languages. It is also relevant to language teachers working in different contexts (e.g., ESL/EFL, K-12 education, higher education) and with learners at different proficiency levels.

This handbook is also unique in three distinct ways. Firstly, it promotes greater collaboration among all stakeholders. This involves encouraging students, parents, teachers, course book writers, editors, and publishers to engage in critical discussions. Such collaborative practices are crucial as Xu et al. (2023) suggest that interactions and negotiations may lead to innovations in materials development. Secondly, it provides a platform for minority voices to be heard. The book includes contributions from a range of experts in the field, providing diverse perspectives on materials development. Many of the chapters are written by experienced practitioners, teachers, or even postgraduate students in materials writing who have yet to be recognized in this field. Lastly, since this book is a compilation of diverse individual perspectives, it enables readers to understand personalized and diversified principles and beliefs. Therefore, it is reasonable to argue that this publication can facilitate joint collaboration, creative discussions, and innovative approaches toward advancing the field of materials development in the future.

Muneera Muftah

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.**

GOBIERNO DE ESPAÑA
MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL

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

https://revistaeducacion.es/