ORIGINAL ARTICLE

ACTFL WILEY
Language Connects

Measuring preservice foreign language teachers' attitudes toward inclusive education through a newly developed scale

Ignacio Fernández-Portero

> **The Challenge**
> Nowadays, it is usual to find students with special educational needs in almost every classroom. Are foreign language teachers ready to teach them? This article provides a newly developed scale to measure their perception of inclusive education, focusing on three dimensions: training, conception of diversity, and support.

University of Extremadura, Cáceres, Spain

Correspondence
Ignacio Fernández-Portero, Teacher Training College, Av. de la Universidad s/n, 10071, Cáceres, Spain.
Email: ignaciofp@unex.es

Abstract

As a consequence of the inclusion of students with special educational needs in mainstream classrooms, there is a need to measure the attitudes of language teachers toward inclusive education. This study used an exploratory sequential mixed methods research design to validate the construct of a newly developed scale. After exploratory qualitative analysis in which 11 experts took part, the instrument was developed and administered to 952 preservice English teachers from Spain and Portugal. The exploratory and confirmatory factor analysis performed confirmed a three-factor structure interpreted as: training, conception of diversity, and support. Results confirm good psychometric properties for the scale and a reliable tool to assess teacher training programs from the perspective of the teaching staff. Moreover, a preliminary analysis of the

survey carried out with the Mann–Whitney U test showed statistically significant differences across countries in eight out of 17 items. Implications for research, policy, and practice are discussed.

KEYWORDS
attitudes, foreign language, inclusion, mixed methods research, scale validation, teacher training

1 | INTRODUCTION

Inclusive education is a complex process in which most of the responsibility falls on teachers, who have to implement the laws and policies through educational practices that benefit all. However, specialist teachers usually feel unprepared for the aforementioned required skills due to a lack of training (Blume et at., 2019; Colmenero Ruiz & Pegalajar Palomino, 2015; Crisol Moya & Caurcel Cara, 2020; Debreli, 2012; Navarro-Mateu et al., 2020; Savic & Prosic-Santovac, 2017; Schwab et al., 2019; Taylor & Ringlaben, 2012). Therefore, if teachers are the key to inclusion (Colmenero Ruiz & Pegalajar Palomino, 2015; Savic & Prosic-Santovac, 2017), they must be trained for such purpose regardless of their specialization.

Little research has studied in-service (ISETs) and preservice English teachers' (PSETs) attitudes toward the inclusion of English language learners (ELLs) with special educational needs (SEN) in mainstream classrooms (Blume et al., 2019; Debreli, 2012; García Montes et al., 2016; Navarro-Mateu et al., 2020; Reeves, 2006; Schwab et al., 2019), as well as the instructional implications of those attitudes (Reeves, 2006) and the students' perception about the use of inclusive practices by their teachers (Navarro-Mateu et al., 2020; Schwab et al., 2019). This type of research tends to focus on academic performance in general and, in fact, none has examined specifically the ways in which SEN students interact with subject-specific issues (Blume et al., 2019). Therefore, more research is needed so that the inclusion of students with disabilities, neurodevelopmental disorders (i.e., attention deficit hyperactivity disorder, autism, etc.), or learning difficulties (e.g., those affecting their thinking, understanding, emotional, behavioral, speech, language, or communication skills) in regular classrooms can be effective (Blume et at., 2019; Debreli, 2012; García et al., 2016; Navarro-Mateu et al., 2020; Reeves, 2006).

The most frequent tools used in the Spanish context are the scales created by González-Gil et al. (2013) and Vélez (2013) focused on preservice teachers, and Colmenero and Pegalajar Palomino (2015) who focused on preservice teachers of secondary education. However, these are generic scales—that is, not centered on a specific subject. The only studies conducted in Spain about PSETs' and ISETs' attitudes toward inclusive education were carried out by Barrios and García (2009), Medina and Pérez (2017), and Crisol Moya and Caurcel Cara (2020). The former was focused on primary education, while the two latter on secondary education. Results from these qualitative studies revealed that both PSETs and ISETs had a positive attitude toward the inclusion of SEN students in mainstream classrooms, although they expressed lack of self-confidence to teach them, lack of training, contradictions between the principles of inclusive education and real classroom practice, lack of resources (both human and material),

lack of coordination and communication between the teachers sharing responsibilities with the same SEN students, and lack of acceptance by non-SEN students.

The problem underlying these studies is that in some of them a measurement instrument has not been developed because they are mainly based on interviews, observation, and/or open-ended questions (see Barrios & García, 2009; Medina & Pérez, 2017), and, when a tool has been developed, it has been designed to be a generic instruments without any single theme, dimension or item related to language education (see Colmenero Ruiz & Pegalajar Palomino, 2015; Crisol Moya & Caurcel Cara, 2020; González-Gil et al., 2013; Vélez, 2013). In the few research focusing on English as an additional language the instruments used have not been shared or details about their psychometric properties have not been provided (see Reoyo et al., 2012; Hernandez & Carrasco, 2012). In Portugal there has not been any study about this subject matter to date and, because of the proximity between these two countries, the numerous exchanges and partnerships among educational institution and centers, and the status of English as a foreign language in education laws, it was considered interesting and practical to have an instrument to assess the quality of language teacher training programs regarding attention to diversity.

In light of the foregoing, the present study seeks to develop and validate the construct of a scale to measure English teachers' attitudes toward inclusive education. The main goal is to create a brief, reliable, and valid instrument that can be easily used and interpreted to help identify how the teaching staff feel about inclusive education and teaching SEN students. Moreover, as a secondary goal, this study aims to analyze the collected data to know whether there are significant differences across countries or not. In the following sections, besides the rationale behind the development of the scale, the qualitative and quantitative methods used will be explained, and the results from the statistical analysis will be described to interpret why the factorial structure proposed is the most appropriate one.

2 | LITERATURE REVIEW

Literature pertaining to teacher training to teach SEN students and the support teachers report receiving to teach them provide context for the creation of the scale in this study.

2.1 | Training

Inclusive EFL is complex and preservice teachers might not feel ready for it when they are about to finish their higher education studies (Blume et al., 2019). For this reason, initial training is essential so that teachers can identify effective teaching strategies if learners are having difficulties (Sharma et al., 2006). In this sense, authorities are responsible for differentiating training programs to ensure they adapt to future teachers' dissimilar needs (Forlin et al., 2009), since the adequate training will improve teachers' attitudes and self-confidence toward inclusion (Forlin et al., 2011).

Negative attitudes are difficult to reverse but they can be changed for positive ones by providing adequate teacher training to EFL teachers (Savic & Prosic-Santovac, 2017). EFL training programs are generally focused on subject matter knowledge and learning about SEN is usually acquired externally. This training, before the teaching practice period, is essential because a major change in the preservice teachers' beliefs appears during this real teaching

experience. When SEN students are involved, PSETs become frustrated and their attitude toward inclusion are negatively affected when they realize that the theory learned cannot be put into practice (Debreli, 2012).

In Spain, ISETs (see Medina & Pérez, 2017) and PSETs (see Reoyo et al., 2012) believe that their training about curriculum-related issues is as important as training about affective and relational skills. They feel underprepared to adapt lessons to the different academic abilities of their students due to lack of strategies and knowledge. Moreover, they also think they need more training about teaching methods and a better connection between theory and practice. For these reasons, EFL teacher training should be considered as the primary tool to ensure SEN students will experience inclusive environments in their EFL classes (Flores & Villardón, 2015; Medina & Pérez, 2017). In the Portuguese context, EFL teachers do not receive training about SEN at university, whereas in Spain, the study of a single theory-based subject about this area is not enough to guarantee that they can successfully meet every student's needs. In this respect, Crisol Moya and Caurcel Cara (2020) carried out the only quantitative study in the context of EFL and are the only researchers proving data about this. In a four-point Likert scale (1 = strongly agree, 4 = strongly disagree) participants ($n = 76$) showed their dissatisfaction with the training received regarding attention to diversity ($M = 2.55$; standard deviation = 0.40).

2.2 | Conception of diversity

Teachers' attitudes can be strongly influenced by the nature and severity of the disabling condition presented to them in the form of child-related, teacher-related, and/or educational environment-related variables (Avramidis & Norwich, 2002). They have prejudices, beliefs, and ideas that act as filters, may affect their decision-making and, therefore, mean an additional barrier to overcome inclusion problems (Debreli, 2012; Sharma et al., 2006). These factors must be considered given that their attitudes toward inclusive schooling influence the individualization of their teaching (Schwab et al., 2019). In fact, there is evidence that teachers who have positive attitudes toward inclusion tend to differentiate more instruction and curriculum during their teaching as it is perceived by their students (Schwab et al., 2019; Taylor & Ringlaben, 2012).

Positive attitudes are predictors of success in inclusive teaching (Savic & Prosic-Santovac, 2017). Since attitudes are not innate but learned, they should be taught to teachers (Navarro-Mateu et al., 2020). For this reason, prior teacher training and updated training programs are essential to prepare them to better work with SEN students, increasing their self-confidence and helping them to develop more positive attitudes toward inclusive practices (Forlin et al., 2011; Navarro-Mateu et al., 2020).

Experience is another factor to consider and, more specifically, from a twofold perspective: first, from direct experience with SEN students, given that previous experiences with them—or lack thereof—nfluence teachers in the way they approach inclusion (Ganschow & Sparks, 2000; Navarro-Mateu et al., 2020). In fact, preservice teachers with more contact with this type of students have more positive attitudes toward them (Belmar et al., 2017). Second, from the experience acquired along the years. Some authors have shown that more experienced teachers tend to have more positive attitudes toward inclusion (Avramidis & Norwich, 2002)—although less experienced teachers (1–5 years of experience) usually implement more academic accommodations due to the preparation of more additional material (Alghazo & Naggar Gaad, 2004; Forlin et al., 2011).

In research conducted by Rodríguez (2005) in a region of southern Spain, it was revealed that preservice primary teachers ($N = 1.404$) had positive attitudes toward inclusion, considering it even a challenge for their professional development, although they admitted they lack knowledge to face such a complex matter. In this study, PSETs ($N = 227$) scored higher than other specialists on items related to the acceptance of SEN students and there was a significant difference regarding their training between first and last-year students. While the former were more positive and thought that they would receive the necessary training to meet every student's need, the latter thought their training was insufficient for such purpose, showing more negative self-efficacy beliefs.

2.3 | Support

As it has been said above, teachers are the key to inclusion, but inclusion without support is difficult to guarantee. This factor must be considered as one of the pillars of inclusive and equitable quality education (UNESCO, 2015), and teachers should be guaranteed not only the tools and resources needed, but also the recognition, labor laws, and professional development to optimize their contribution (UNESCO, 1990).

Inclusion must be carefully planned and well supported so that teachers' initial reservations or concerns are overcome (Avramidis & Norwich, 2002). Governments and administrations, therefore, play a crucial role since the legal framework and resource management in which all formal education is encompassed, entirely depend on them. Their inaction can cause indifference and demotivation among the teaching staff (García et al., 2016) and this cannot be afforded by SEN students. As an example, Savic and Prosic-Santovac (2017) revealed that 87.5% of the EFL teachers polled in their research showed negative attitudes toward inclusion due to the absence of adequate conditions for effective inclusive practice. Hence, to change negative for positive attitudes, authorities should improve the conditions in our schools (Savic & Prosic-Santovac, 2017), make the allocation of the available resources more flexible (based on the severity of needs represented in the inclusive settings) and set learning support teams within the schools to care for individual teachers who request guidance over a teaching concern related to SEN (Avramidis & Norwich, 2002).

So that this teamwork can success, certain collaboration between teachers and any other professional working with SEN students must be required. Without teachers' coordination curricula and instruction could not be properly adapted in every subject to meet every student's needs (UNESCO, 1994). Unfortunately, collaboration can be improved most of the time (Colmenero Ruiz & Pegalajar Palomino, 2015; Schwab et al., 2019). Savic and Prosic-Santovac (2017) reported that almost 70% of respondents in their study carried out in Serbia admitted that SEN teams rarely met, and in Spain, Reoyo et al. (2012) and Hernández and Carrasco (2012) showed that preservice secondary teachers think that the lack of coordination and communication perceived during their practicum makes total inclusion hard to achieve.

3 | METHOD

An exploratory sequential mixed methods research design was used to develop the English Teachers' Attitudes Toward Inclusive Education Scale (ETATIES). This method employs "rigorous quantitative research assessing magnitude and frequency of constructs and rigorous

qualitative research exploring the meaning and understanding of constructs" (Creswell et al., 2011, p. 4). To do so, the guidelines, psychometric standards, and criteria explained in the following sections were used to assess the test's appropriateness, meaningfulness, and usefulness. These qualities were mainly determined through factor analysis (see Brown, 2014; Brown & Moore, 2015; Comrey & Lee, 1992) and validity/reliability tests (see Benson & Clark, 1982; Cook & Beckman, 2006; Nunnally & Bernstein, 1994).

The research design followed a sequence in which one phase of collection and/or analysis was followed by another. First, the first draft of the scale was developed—according to similar research and the extant literature—and sent to a group of experts. Second, the experts' feedback was collected and the first draft was modified according to their comments and assessment, which led to the hypothesis of a possible three-factor solution as the most appropriate model for this scale. Thirdly, the first test was carried out with 322 participants to assess the factorial structure of the scale. Once the statistical analysis was performed, the corresponding adjustments were made to develop the final questionnaire, which was filled in by 952 participants. Finally, a factor analysis with the quantitative data collected was run. This research design is shown in the diagram of Figure 1.

3.1 | Qualitative phase

A qualitative approach facilitates the collection of data when measures do not exist and provide a depth of understanding of concepts (Creswell et al., 2011). This understanding was the result of the analysis of the participants' opinions. The process is explained in the following sections.

3.1.1 | Participants

Once a thorough literature review was carried out about scales measuring teachers' attitudes toward the inclusion of SEN students in regular classrooms, a first draft of the scale was sent to
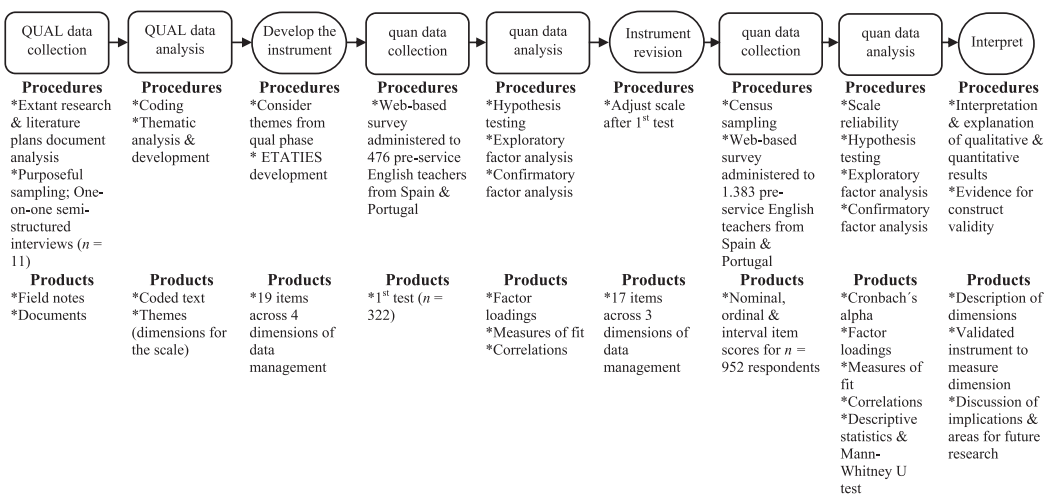

FIGURE 1 Sequential exploratory mixed methods design diagram

20 experts in the field, out of whom 11 accepted to participate in the development and content validity of this new instrument. To identify them, the purposeful sampling technique developed by Patton (2002) was followed. The aim was to select information-rich individuals who could yield insights and in-depth understanding rather than empirical generalizations through their knowledge and experience about the topic (Creswell & Plano Clark, 2011; Patton, 2002). By looking for demographic variation, different points of view were explored from the two main areas covered in this research: bilingual and inclusive education. The 11 experts who finally took part in this exploratory phase hold a PhD and their fields of research are related to Foreign/English Language Teaching or Educational Psychology. Seven of them were Spanish, three Portuguese, and one was a US citizen. They have a broad international research experience and know both the Portuguese and the Spanish education systems. Table 1 below shows more detailed information about these participants.

3.1.2 | Data collection

As part of the methodological structure selected, this study was aimed to have an overall qualitative approach with data collection through expert interviews with a purposeful sample, which would lead to determine key terms to then display findings when building scenarios, as Muskat et al. (2012) suggest. Therefore, these experts were interviewed using the semi-structured protocol proposed by Holstein and Gubrium (1995), besides researcher notes. Once the first draft was sent (a 28-item scale), their comments and suggestions were classified into different categories (theme-related, item-related, etc.) to improve the questionnaire. The purpose of this stage was threefold: (a) to correct any mistake made in the scale's content (items, dimensions, wording...), (b) to avoid ambiguity, bias, or any type of persuasiveness that could influence the participants' answers, and (c) to adjust the scale to the education system of the two countries it was aimed at: Spain and Portugal.

3.1.3 | Data analysis

NVivo software was used for qualitative data management and analysis, which allowed to classify, sort, and arrange the information collected in the interviews. Relationships in the data were then analyzed and in vivo codes were created from terms and phrases from the interview transcript (Muskat et al., 2012). Recurring patterns were detected from these codes to finally generate a smaller number of themes (Charmaz, 2006). As a result of this analysis, the first draft was modified. The initial 28 items and six themes were reduced to 19 and three, respectively. Moreover, judgments were aimed to be collected using a five-point Likert scale (1 = strongly disagree; 5 = strongly agree) to allow for higher reliability in capturing participants' opinions and because it usually increases not only response rates, but also response quality due to their less confusing structure (Babakus & Mangold, 1992; McKelvie, 1978). During this modification process, 12 items remained untouched, nine were removed, and seven were rewritten.

The final themes and categories were established once they were internally homogeneous and externally heterogeneous, as Patton (2002) suggests. Categorical codes helped to separate condense data into categories so that the main themes could be captured and synthesized

TABLE 1 Information about the participants for the exploratory qualitative phase

Nationality	Sex	Credentials	University	Current position
Portuguese/Spanish	F	PhD in English Language Teaching	University of Extremadura	Lecturer
Spanish	F	PhD in English Language Teaching	University of Extremadura	Professor of English Philology
Spanish	M	PhD in English Language Teaching	University of Alcalá	Senior Lecturer in English Philology
Spanish	F	PhD in Educational Psychology	University of Extremadura	Lecturer
Spanish	M	PhD in Educational Psychology	Autonomous University of Barcelona	Lecturer
Spanish	F	PhD in English Language Teaching	Autonomous University of Barcelona	Professor of Language Education
Spanish	F	PhD in English Language Teaching	University of Granada	Lecturer
Spanish	F	PhD in Educational Psychology	University of Extremadura	Lecturer
US	F	PhD in Hispanic Language and Literature	University of North Carolina	Professor of Spanish
Portuguese	F	PhD in English Studies	Castelo Branco Polytechnic Institute	Professor of English Cultural Studies
Portuguese	F	PhD in Bilingual Education	University of Porto	Lecturer

(Charmaz, 2006). Such categories finally explained ideas, concepts, and processes, which led to the generation of a precise wording of the items.

The process described above resulted in three themes: training, conception of diversity, and support. The first theme aims to cover the most basic aspects related to EFL teacher training: the need of a specific training to teach English to SEN students, the level of satisfaction with the training received, and concepts related to preservice teachers' capacity to adapt the core concepts of education programs (such as the methods, contents, materials, and assessments) to every student's needs. The second theme focuses on the teaching of SEN students in regular classrooms, highlighting aspects such as expectations, acceptance, or workload. Finally, the third theme indicates some of the support available for both teachers and SEN students in the educational context, such as resources, teacher collaboration, regulations, and information about special educational needs to raise public awareness.

3.1.4 | Validity and reliability

To improve trustworthiness and internal validity, triangulation, member checking, and detailed reporting (Creswell, 2007) were included in the method. Triangulation was used to find common trends from the experts and bring subjects together (Patton, 2002). This process was repeated several times to reduce researcher bias in determining key terms while adding credibility. Realistic conclusions were drawn from member checking and detailed reporting to find the underlying meaning of the experts' opinions. The goal was to look for rigor and validity by sending them the description of the process leading to the final version of the scale and asking for their approval—that is, using the synthesized and analyzed data method proposed by Harvey (2015). This was essential to ensure that their own meanings and perspectives were represented in the instrument to be used (Tong et al., 2007). The purpose of the entire process was twofold: to ensure that the results truthfully reflected the facts studied and that similar results would be obtained if the study were replicated by others.

3.2 | Quantitative phase

The qualitative phase implies a solid scaffold for the quantitative part of the study, which will be explained in this section. Data from the sample will be facilitated as well as the description of the procedures that were followed to obtain the statistical results.

3.2.1 | Sample

Participants in this study had to meet the following two requirements:

- To be last-year students specialized in EFL teaching in the following degrees: Bachelor of Arts (BA) in Preschool Education, BA in Primary Education, Master of Arts (MA) in EFL teaching in Secondary Education (in the case of preservice Spanish teachers), or MA in EFL teaching in Primary and/or Secondary Education (in the case of preservice Portuguese teachers).
- To have carried out at least one practicum.

TABLE 2 Data from the sample: participants per country, age, gender, and universities taking part in the study.

	n	Age	Gender M	F	Universities eligible for the study	Universities taking part in the study
Spain	912	22.7 (SD = 2.78)	253 (27.7%)	659 (72.3%)	71 (out of 87)	30
Portugal	40	26.5 (SD = 4.99)	3 (7.5%)	37 (92.5%)	7 (out of 37)	4
Total	952	22.89 (SD = 3.0)	256	696	78 (out of 124)	34

Abbreviation: SD, standard deviation.

To find all the higher education centers where these degrees were offered, a thorough search was made through the academic divisions' websites of each university in these two countries. Since in most Spanish universities the educational offer regarding the BAs mentioned above is bilingual and/or non-bilingual, each syllabus was analyzed to know whether they were considered as such according to the National Quality Assessment Body or not. Table 2 shows the data from the sample, and it can be noticed two significant aspects: there are outstandingly (a) more women than men in this profession and (b) more Spanish universities than Portuguese ones.

Once the universities were identified, an e-mail was sent to the teaching staff in charge of last-year students who were carrying out a practicum and met the inclusion criteria. In that e-mail the research was briefly introduced, and the teaching staff were asked to disseminate the online survey form among their students. For the following step, personalized follow-up reminder e-mails were sent until the set deadline.

3.2.2 | Response rate

According to the information gathered in the Google Forms database, there were two high peaks where most responses were received (61%). Those two peaks took place during the first and third week of June, respectively. Finally, the total number of respondents was 952 out of 1383, as reported by the contacts from every faculty, which means a 68% of response rate. Therefore, the sample size is representative, since 301 were enough considering (a) ± 5 margin of error, (b) 50% heterogeneity, and (c) 95% confidence intervals.

3.2.3 | Data analysis

The set of data collected from the participants was first assessed using exploratory factor analysis (EFA) with principal axis factoring and orthogonal varimax rotation since it was aimed to identify the underlying relationships between the measured variables and latent constructs. Then, polychoric correlations were run given that variables and data were ordinal. Bartlett's, (1950, 1951) test and Kaiser-Meyer-Olkin's (KMO) test were used to measure sample adequacy and whether it was appropriate to perform a factor analysis or not. Once the dimensionality of the scale was assessed, the Kaiser's (1960) eigenvalue criterion (i.e., greater than 1.0) and a scree plot led to establish the number of factors (Cattell, 1966).

Once the EFA was conducted, a confirmatory factor analysis (CFA) was performed to test the proposed factor model structure (Brown, 2014). The model fit was then evaluated with structural equation modeling using goodness of fit measures such as (a) χ^2 statistics, (b) comparative fit index (CFI), (c) Tucker–Lewis index (TLI), (d) Normed Fit Index (NFI), (e) standardized root mean square residual (SRMR), (f) root mean square error of approximation (RMSEA), and (g) Akaike Information Criterion (AIC). Cronbach's α was used to measure the overall internal consistency and how closely related the items were as a group (Cronbach, 1951). The final survey consisted of 17 items distributed in three dimensions (teaching, conception of diversity, and support) and demographic items.

4 | RESULTS

EFA and CFA were carried out to measure the participants' attitudes and perceptions about their training in higher education, the support available in schools, and their conception of diversity (in terms of including SEN students in regular classrooms). For the first model, a 19-item scale was filled in by 322 participants. The results from factor analyses led the research to a slight modification of the scale and to hypothesize a better model fit for the resultant questionnaire. For the second model, the 17-item scale had 952 responses and showed a better model fit.

4.1 | Preliminary data analysis

The KMO test was used to predict if data were likely to factor well based on correlation and partial correlation (Hutcheson & Sofroniou, 1999). The KMO measure was 0.90 for Model 1, which indicated that the sample adequacy was "superb" (Hutcheson & Sofroniou, 1999, p. 225), and Bartlett's test was statistically significant ($p < .01$). Then, since the five-point Likert scale was ordinal, polychoric correlations were estimated to measure the correlations between items. As a result, two items were removed from Model 1 to increase the test reliability because they had low factor loadings. In Model 2 all the items displayed a correlation between .67 and .90, and KMO = 0.92.

4.2 | Exploratory factor analysis

Among the existing methods to determine the number of factors to retain, Bartlett's test, Kaiser's (1960) eigenvalue criterion, and Cattell's (1966) scree test were used. In Model 1, the scree plot suggested the presence of four factors, which explained 69.1% of the variance. However, after two items were removed due to their low correlations, the three-factor solution fitted the data better than the previous model, according to the EFA performed. This factor analysis was based on polychoric correlations and it was carried out with unweighted least squares and varimax rotation on the set of ordinal data.

4.3 | Retaining items

The appropriate estimation of the number of factors to retain when developing a new scale is of significance for its successful application (Courtney & Gordon, 2013). Therefore, it is a

TABLE 3 Factor loadings from exploratory analysis in Model 1

	Component			
	1	2	3	4
C.2	0.88			
C.4	0.87			
C.1	0.88			
C.3	0.87			
C.5	0.86			
C.8	0.67			
C.6	0.76			
C.7	0.69			
T.3		0.84		
T.2		0.86		
T.4		0.84		
T.1		0.90		
T.5		0.74		
T.6		0.56		0.48
S.2			0.88	
S.4			0.80	
S.3			0.83	
S.1			0.62	0.51
S.5			0.71	

Note: Factor extraction method: Principal Components Analysis. Rotation method: Varimax with Kaiser normalization. Blanks represent < 0.40.

procedure that must be carefully carried out so that under- and over-extraction do not alter the interpretation of EFA results in the data reduction process.

The three methods mentioned in the previous section contributed to determine the number of factors to retain with the ordinal data collected from Model 1. In this 19-item scale, a factor analysis was performed with polychoric correlations, using a principal component analysis (PCA) and the mean eigenvalue criterion as the data extraction method. The two items removed belonged to the themes *training* (T6) and *support* (S1). They loaded 0.56 and 0.62, respectively, which can be acceptable. However, they loaded poorly on the fourth factor and data suggested that removing these two items would increase reliability. Table 3 below shows the factor loadings of Model 1. These changes later proved to be successful since the EFA from Model 2 showed better factor loadings for each item and the subsequent three-factor solution could explain 71.1% of the variance. Moreover, the CFA stated below confirmed that the removal of these two items improved the model fit.

4.4 | Confirmatory factor analysis results

EFA was run with the tools available in SPSS, while the CFA was implemented in IBM® SPSS® Amos by using the maximum likelihood estimator. This multivariate statistical procedure was conducted with both models to confirm that the second one had a better factor structure.

4.5 | Model fit

The first model, based on four factors, showed a poor fit to the data (χ^2 (118) = 793.575, χ^2/df = 1.53, $p < .001$, CFI = 0.868, TLI = 0.829; NFI = 0.850; SRMR = 0.079; RMSEA = 0.075; AIC = 897.575; PRATIO = 0.771). Then, to improve the model fit, a few changes were made by using standardized residuals and modification indices. These two statistics helped in the identification of two areas of misfit in the CFA solution, which led to the elimination of one item (T6) due to high standardized residuals (≥1.96; $p < .05$; Agresti, 2002) and another one (S1) due to a low factor loading (<0.45; Comrey & Lee, 1992). For the next and last adjustment, modification indices above 20 were reviewed, which led to the relation of three pairs of measured variables (T2 and T4, C2 and C4, C6 and C7). Correlated measurement error was then modeled in the CFA solution on the basis of method effects due to possible differential susceptibility (such as stereotypes or preconceived ideas shared by the participants on these six items). Once these modifications were made, in the following CFA both CFI and TLI were above 0.95. Table 4 shows the item-to-factor statistics from the CFA of Model 1 and the justification of the removal of these two items.

After the removal of these two items, the result was a less balanced scale in terms of item distribution (1st dimension = 5 items, 2nd dimension = 8 items, 3rd dimension = 4 items), although these changes neither affected the scale conceptually nor theoretically, and the model fit was improved showing the following results related to Model 2: χ^2 (116) = 362.146, χ^2/df = 1.36, $p < .001$; CFI = 0.979; TLI = 0.975; NFI = 0.969; SRMR = 0.05; RMSEA = 0.047; AIC = 470.146; PRATIO = 0.853.

4.6 | Internal consistency

Reliability and internal consistency were assessed through Cronbach's α coefficient and were considered optimal (Nunnally & Bernstein, 1994) since the overall reliability was .88. Coefficients on the three factors were as follows:

- Factor 1 (training → T1–T5) coefficient at .89
- Factor 2 (conception of diversity → C1–C8) coefficient at .93
- Factor 3 (support → S2–S5) coefficient at .85

4.7 | Summary of results: Loadings, correlations, and reliability

The main goal of this study was to create and validate a scale to measure preservice English teachers' attitudes toward inclusive education. Exploratory methods carried out on the collected data ($n = 952$) contributed to interpret a three-factor model based on three

TABLE 4 Item-to-factor statistics from confirmatory factor analysis

Item	Factor Loading	Standard error	Residual	R^2	MI > 20	Removed
T1	0.87	0.02	0.96	0.58		
T2	0.82	0.03	0.52	0.45		
T3	0.77	0.04	0.63	0.26		
T4	0.86	0.04	0.94	0.38		
T5	0.72	0.03	0.45	0.76		
T6	0.54	0.06	2.09	0.22	●	●
C1	0.86	0.03	0.69	0.74		
C2	0.85	0.03	0.35	0.29		
C3	0.87	0.04	0.39	0.67		
C4	0.82	0.03	0.32	0.27		
C5	0.80	0.04	0.25	0.53		
C6	0.68	0.03	0.34	0.26		
C7	0.75	0.02	0.43	0.72		
C8	0.63	0.04	0.42	0.76		
S1	0.42	0.05	1.52	0.37		●
S2	0.83	0.02	0.16	0.43		
S3	0.78	0.03	1.18	0.65		
S4	0.75	0.03	1.21	0.72		
S5	0.66	0.02	0.97	0.34		

dimensions: (a) teaching, (b) conception of diversity, and (c) support. Then, a maximum likelihood estimation was performed with a CFA to test the factor structure. The analysis concluded that a 17-item scale on three factors had the best model fit. Internal consistency and reliability were considered optimal through Cronbach's alpha and goodness of fit criteria. The result (see Table 5) is based on an empirical model designed to measure future English teachers' attitudes toward inclusive education.

4.8 | Preliminary analysis of survey results

A preliminary analysis of the survey results is showed in Table 6. Differences across countries on means and standard deviations reveal considerable differences in items T5, C2, C4, C6, S2, and S4. Additionally, since the rating scale used yielded ordinal level measures, the non-parametric Mann–Whitney U test was used to compare differences between the two independent groups (Spanish and Portuguese respondents). The test showed statistically significant differences in items T1, T5, C2, C4, C6, S2, S4, and S5, finding the biggest differences in items C2, C4, C6, S2, and S4.

TABLE 5 ETATIES

Item	Factor loading	Item-test correlation	Subscale reliability
Training			0.89
T1. Inclusion of SEN students in regular classes requires a specific training for English language teachers focused on special educational needs.	0.87	0.86	
T2. I think more training is necessary to know how to adapt the teaching methods to SEN students.	0.82	0.73	
T3. I think more training is necessary to know how to adapt the teaching materials to SEN students.	0.77	0.73	
T4. I think more training is necessary to know how to adapt assessment tools to SEN students.	0.86	0.68	
T5. I am satisfied with the training I have received to teach every student in my EFL class.	0.74	0.64	
Conception of diversity			0.93
C1. SEN students cannot learn English as a foreign language up to the same standard as non-SEN students.	0.86	0.76	
C2. I think it will be difficult to provide non-SEN students with appropriate academic challenges in an inclusive English classroom.	0.85	0.81	
C3. Having SEN students in my EFL class would increase my workload.	0.87	0.82	
C4. Having a student with an individualized academic program in my EFL class will have negative effects in the classroom environment.	0.82	0.81	
C5. SEN students are usually accepted by their peers	0.80	0.80	
C6. Inclusion of SEN students in EFL classes can benefit non-SEN students.	0.68	0.70	
C7. Inclusion facilitates socially appropriate behavior among all students.	0.75	0.60	
C8. SEN students can develop their academic skills more rapidly in regular classrooms.	0.64	0.49	
Support			0.85
S2. There are insufficient resources available to support SEN students in an inclusive English classroom.	0.83	0.59	
S3. Having more opportunities to collaborate with other teachers in the classroom would enrich the teaching-learning processes.	0.78	0.75	
S4. Current curricula create more barriers than support for SEN students.	0.75	0.68	

TABLE 5 (Continued)

Item	Factor loading	Item-test correlation	Subscale reliability
S5. More information about SEN students (their skills and needs) among all the members of the educational community (teachers, parents, students, school board...) would benefit their inclusion.	0.66	0.60	

Note: χ^2 (116) = 362.146, χ^2/df = 1.36, $p < .001$, CFI = 0.979, TLI = 0.975, NFI = 0.969, SRMR = 0.05, RMSEA = 0.047, AIC = 470.146, PRATIO = 0.853, α = .88.

Abbreviations: CFI, comparative fit index; ETATIES: English Teachers' Attitudes Toward Inclusive Education Scale; RMSEA, root mean square error of approximation; SEN, special educational needs; SRMR, standardized root mean square residual; TLI, Tucker–Lewis index.

TABLE 6 Response means, SD, and Mann–Whitney U test for ETATIES.

Item	Spanish ($n = 912$)		Portuguese ($n = 40$)		Mann–Whitney U test	
	M	SD	M	SD	U	p Value (<.05)
T1	4.63	0.523	4.40	0.545	14238.000	.005
T2	4.68	0.517	4.65	0.533	17650.000	.662
T3	4.64	0.544	4.65	0.533	18148.000	.947
T4	4.52	0.669	4.63	0.540	17125.500	.447
T5	2.19	0.859	1.88	0.648	14681.000	.027
C1	3.07	0.934	2.80	1.018	15074.500	.051
C2	3.21	0.967	2.70	0.823	12538.500	.000
C3	4.46	0.641	4.35	0.662	16486.500	.245
C4	3.16	0935	2.25	0.630	8170.000	.000
C5	3.71	0.750	3.90	0.441	15898.000	.120
C6	2.73	0.988	3.60	0.810	9515.000	.000
C7	4.27	0.624	4.35	0.533	17210.500	.496
C8	3.32	0.976	3.58	0.874	15429.500	.082
S2	4.40	0.702	4.05	0.597	12553.000	.000
S3	4.59	0.518	4.48	0.554	16318.500	.186
S4	4.20	0.807	3.73	0.679	11794.500	.000
S5	4.66	0.505	4.48	0.554	15032.000	.021

Abbreviation: SD, standard deviation.

5 | DISCUSSION

The purpose of this study was twofold. In the first place, this research was aimed at creating and validating the construct of a scale through which preservice English teachers could express their attitudes and perceptions about three dimensions concerning inclusive education: the

training received during their BA or MA, their conception of diversity, and the support available in education centers. And, in the second place, this study aimed to analyze the collected data so that the results could contribute to future research and have practical implications for teacher training in terms of improving educational policies and teaching practice.

The factor structure of the ETATIES can be considered as largely confirmed and suggests that this newly developed scale is full of promise as a valid and reliable measure of teachers' attitudes toward inclusion for research, policy, and practice. The PCA carried out revealed a three-factor structure and individual analysis of items suggests that the 17-item scale distributed across the aforementioned three dimensions was the most appropriate statistically. Table 7 illustrates the item distribution among the three factors in the final scale. The reliability test of the tool indicates high internal consistency (Cronbach's $\alpha = .88$), which provides credibility for the developed instrument.

The items were chosen on an empirical basis, considering similar scales, and trying to develop a balanced tool while making it as short as possible to encourage participants to respond to the whole instrument (demographic data included). Finally, in terms of item distribution, the scale is not as balanced as it was intended, but the final result is the most recommended statistically and, therefore, the most reliable one. In terms of adaptability for its implementation, ETATIES can be easily adapted to other content areas simply by changing the subject under study in items 1, 6, 7, and 14.

Regarding the preliminary analysis of the survey results, the means from the descriptive statistics show that both Portuguese and Spanish PSET have similar attitudes toward their training, feel uncertain about SEN students being able to learn English as a foreign language up to the same standards as non-SEN students (Portuguese students feel slightly more pessimistic than Spanish ones), and agree when considering that: (a) having SEN students would increase their workload, (b) SEN students are usually accepted by their peers, (c) inclusion facilitates socially appropriate behavior among all students, (d) SEN students can develop their academic skills more rapidly in regular classrooms, and (e) having more opportunities to collaborate with other teachers in the classroom would enrich the teaching-learning processes. However, the Mann–Whitney U test showed significant differences in eight out of 17 items. In these cases, Portuguese students showed more dissatisfaction with the training received to address every student's needs and believe that the inclusion of SEN students in regular classes requires a specific training for English teachers to a lesser degree than Spanish students. When assessing their conception of diversity, Portuguese participants showed more positive attitudes toward SEN students because they think that: (a) it would not be so difficult to provide non-SEN students with appropriate academic challenges in an inclusive English classroom, (b) having students with individualized academic programs in EFL classes does not have to negatively affect the classroom environment, and (c) their inclusion can benefit non-SEN students. Moreover, in terms of support, Portuguese PSET, alike Spanish ones, think that the resources available to support SEN students in inclusive English classroom are insufficient and more information about them (e.g., their skills and needs) among all the members of the educational community would benefit their inclusion, although to a lesser degree than Spanish PSET in these two items. Finally, Portuguese students have a more positive conception of curricula since they do not think it is such a barrier as Spanish students think.

TABLE 7 ETATIES: Three dimensions and corresponding items

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Training						
1	Inclusion of SEN students in regular classes requires a specific training for English language teachers focused on special educational needs.	1	2	3	4	5
2	I think more training is necessary to know how to adapt the teaching methods to SEN students.	1	2	3	4	5
3	I think more training is necessary to know how to adapt the teaching materials to SEN students.	1	2	3	4	5
4	I think more training is necessary to know how to adapt assessment tools to SEN students.	1	2	3	4	5
5	I am satisfied with the training I have received to teach every student in my EFL class.	1	2	3	4	5
Conception of diversity						
6	SEN students cannot learn English as a foreign language up to the same standards as non-SEN students.	1	2	3	4	5
7	I think it will be difficult to provide non-SEN students with appropriate academic challenges in an inclusive English classroom.	1	2	3	4	5
8	Having SEN students in my EFL class would increase my workload.	1	2	3	4	5
9	Having a student with an individualized academic program in my EFL class will have negative effects in the classroom environment.	1	2	3	4	5
10	SEN students are usually accepted by their peers	1	2	3	4	5
11	Inclusion of SEN students in EFL classes can benefit non-SEN students.	1	2	3	4	5
12	Inclusion facilitates socially appropriate behavior among all students.	1	2	3	4	5

(Continues)

TABLE 7 (Continued)

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
13 SEN students can develop their academic skills more rapidly in regular classrooms.	1	2	3	4	5
Support					
14 There are insufficient resources available to support SEN students in an inclusive English classroom.	1	2	3	4	5
15 Having more opportunities to collaborate with other teachers in the classroom would enrich the teaching-learning processes.	1	2	3	4	5
16 Current curricula create more barriers than support for SEN students.	1	2	3	4	5
17 More information about SEN students (their skills and needs) among all the members of the educational community (teachers, parents, students, school board...) would benefit their inclusion.	1	2	3	4	5

Abbreviations: ETATTIES: English Teachers' Attitudes Toward Inclusive Education Scale; SEN, special educational needs.

5.1 | Implications for research, policy, and practice

This study has shown potential for ETATIES to be a useful instrument to measure PSETs′ attitudes toward inclusive education from the perspectives of assessing their training, their conception of diversity, and the support available in schools. However, despite the positive results obtained, this research has some limitations. In this section, limitations and implications for future research, education policy, and classroom practice will be discussed.

5.1.1 | Limitations

The first limitation relates to the items and dimensions. The ones showed here are the result of the convergence between qualitative and quantitative research, although other areas of English teachers′ attitudes toward inclusive education can be explored. The second limitation entails internal validity and reliability since this scale is a self-report survey. In self-report questionnaires the phenomenon known as *social desirability bias* can make respondents not to answer truthfully, especially on sensitive questions, and respond in a socially acceptable way regardless of the question (Demetriou et al., 2014). In any case, continued validity testing may be a future research direction. The third and last limitation concerns the sample, given that the instrument was only validated in the Spanish and Portuguese context and it may not be applicable in other contexts, or it may need to be adapted.

5.1.2 | Implications for policy

In a time when results are sometimes more important than skills and knowledge, little time and effort is devoted to guarantee inclusion (UNESCO, 1994, 2015). If teachers are not well trained to address every student's needs, inequalities will exist. For this reason, the provision of extensive opportunities for training at pre- and in-service levels should be seen as a top priority for policy-makers (Avramidis & Norwich, 2002). In this regard, this study provides a potential tool to assess PSETs′ perceptions on the construct. As long as future reliability and validity testing is carried out, this scale offers an instrument that can be used to update and improve training programs at university level, which should be constantly reviewed and adapted to the new social and cultural scenarios (Colmenero Ruiz & Pegalajar Palomino, 2015; Crisol Moya & Caurcel Cara, 2020). Inclusive education has a short history in the two countries under study (the most significant education reforms date from 2018 in Portugal and 2013 in Spain: Law 54/ 2018 in the case of the former and Law 8/2013 in the latter) and it can be seen as a barrier that teachers are not ready to overcome. If they receive not only the training but also the assistance needed, they will become more committed and effective to face obstacles as their effort and skill increase (Avramidis & Norwich, 2002).

5.1.3 | Implications for practice

Adequate training, along with knowledge and expertize must lead to the expansion of educational opportunities and total inclusion. Teachers must be able to implement the educational policies in which all scenarios must be considered. For this to happen, policies and

curricula must be flexibly designed and, if so, they will meet more learners' needs and facilitate that an entire class can progress at a similar pace. Another essential element that can lead to success is coordination between teachers sharing responsibilities with the same SEN students, which must be permanent so that classroom accommodations are effective for everyone. Finally, awareness can be increased and lead to more positive attitudes and better acceptance toward SEN students if information about them (weaknesses, strengths, learning styles, etc.) is disseminated in educational contexts (UNESCO, 1994, 2015). The combination of all these actions should provide SEN students with suitable assistance and guidance so that they can develop their skills and achieve the learning goals.

6 | CONCLUSION

The main goal for the ETATIES was to develop and validate a construct that could represent the current situation of inclusive education in Spain and Portugal. For such purpose, an exploratory sequential mixed methods research design was used, in which a series of qualitative and quantitative procedures were utilized for data collection and analysis. As a result, three dimensions and a 17-item scale can be employed to identify the weaknesses and strengths of both education systems and teacher training programs. As a secondary goal, this study aimed to analyze the collected data to know whether there were significant differences across countries or not. Findings suggest that both Spanish and Portuguese PSET feel disappointed about their training and the support available in schools in a similar way, but Portuguese students feel more optimistic about having SEN students in regular classrooms and do not feel curricula are such a barrier as Spanish students think.

This project, aimed to justify the need to create this scale, describes the development and validation of the instrument so that PSETs could express their perception about the inclusion of SEN students in mainstream classrooms, and examine its potential contribution to improving training programs. Extant literature suggests that well trained teachers are more willing to implement effective education practices, which are essential to provide students with equal learning opportunities in an inclusive and sustainable education system regardless of their educational need. Ongoing research within this domain is of particular importance so that English or language teachers are qualified to face the challenges in current language education.

ORCID
Ignacio Fernández-Portero http://orcid.org/0000-0002-3323-217X

REFERENCES
Agresti, A. (2002). *Categorical Data Analysis* (2nd ed.). Wiley.
Alghazo, E. M., & Naggar Gaad, E. E. (2004). General education teachers in the United Arab Emirates and their acceptance of the inclusion of students with disabilities. *British Journal of Special Education*, *31*, 94-99. https://doi.org/10.1111/j.0952-3383.2004.00335.x

Avramidis, E., & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: A review of the literature. *European Journal of Special Needs Education*, *17*(2), 129-147. https://doi.org/10.1080/08856250210129056

Babakus, E., & Mangold, W. G. (1992). Adapting the SERVQUAL scale to hospital services: An empirical investigation. *Health Services Research*, *26*, 767-786.

Barrios, M. E., & García, J. (2009). Barriers to integrative education from the perspective of pre-service primary teachers of English. *RELIEVE*, *15*(1). https://ojs.uv.es/index.php/RELIEVE/article/view/4185/3796

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*, 77-85.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Statistical Psychology*, *4*, 1-2. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Belmar, M., Cornejo, J., Contreras, P., & Mellado, C. (2017). Modelo factorial de inclusión en estudiantes de pedagogía. *Cuadernos de Investigación Educativa 8*(2), 49-65. https://doi.org/10.18861/cied.2017.8.2.2684

Benson, J., & Clark, F. (1982). A guide for instrument development and validation. *The American Journal of Occupational Therapy*, *36*(12), 789-800. https://doi.org/10.5014/ajot.36.12.789

Blume, C., Gerlach, D., Roters, B., & Schmidt, T. (2019). The ABCs of inclusive English teacher education: A quantitative and qualitative study examining the attitudes, beliefs and (reflective) competence of pre-service foreign language teachers. *The Electronic Journal for English as a Second Language*, *22*(4), 1-18.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Publications.

Brown, T. A., & Moore, M. T. (2015). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 361-379). Guilford Press.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Charmaz, K. (2006). *Constructing Grounded Theory. A Practical Guide through Qualitative Analysis*. Sage.

Colmenero Ruiz, J., & Pegalajar Palomino, C. (2015). Questionnaire for future teachers of secondary perceptions about attention to diversity: Design and validation of the instrument. *Estudios sobre Educación*, *29*, 165-189. https://doi.org/10.15581/004.29.165-189

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum.

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166.e7-166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036

Courtney, M., & Gordon, R. (2013). Determining the number of factors to retain in EFA: using the SPSS R-menu v2 0 to make more judicious estimations. *Practical Assessment,Research, and Evaluation*, *18*(1):8. https://doi.org/10.7275/9cf5-2m72

Creswell, J. W. (2007). *Research design: Qualitative and quantitative approaches* (2nd ed.). Sage.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). *Best practices for mixed methods research in the health sciences*. National Institutes of Health.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage.

Crisol Moya, E., & Caurcel Cara, M. J. (2020). Percepciones de los estudiantes de la especialidad de lengua extranjera - inglés sobre atención a la diversidad en la formación inicial del profesorado de Educación Secundaria. *Onomázein Revista de lingüística filología y traducción*, *6*, 167-188. https://doi.org/10.7764/onomazein.ne6.09

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334. https://doi.org/10.1007/BF02310555

Debreli, E. (2012). Change in beliefs of pre-service teachers about teaching and learning English as a foreign language throughout an undergraduate pre-service teacher training program. *Procedia-Social and Behavioral Sciences*, *46*, 367-373. https://doi.org/10.1016/j.sbspro.2012.05.124

Demetriou, C., Ozer, B. U., & Essau, C. A. (2014). Self-report questionnaires. In R. L. Cautin, & S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 1-6). John Wiley & Sons.

Flores, L., & Villardón, L. (2015). Actitudes hacia la inclusión educativa de futuros maestros de inglés. *Revista Latinoamericana de Educación Inclusiva*, *9*(1), 63-75.

Forlin, C., Earle, C., Loreman, T., & Sharma, U. (2011). The sentiments, attitudes, and concerns about inclusive education revised (SACIE-R) scale for measuring pre-service teachers' perceptions about inclusion. *Exceptionality Education International*, *21*(3), 50-65. https://doi.org/10.5206/eei.v21i3.7682

Forlin, C., Loreman, T., Sharma, U., & Earle, C. (2009). Demographic differences in changing pre-service teachers' attitudes sentiments and concerns about inclusive education. *International Journal of Inclusive Education*, *13*(2), 195-209. https://doi.org/10.1080/13603110701365356

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50. https://doi.org/10.2307/3151312

Ganschow, L., & Sparks, R. L. (2000). Reflections on foreign language study for students with language learning problems: research, issues and challenges. *Dyslexia*, *6*, 87-100. https://doi.org/10.1002/(SICI)1099-0909(200004/06)6:2<87::AID-DYS153>3.0.CO;2-H

García Montes, P. A., Nieto Caballero, J., & Arbeláez Rodríguez, C. (2016). English teachers' attitudes toward inclusive education: A literature review. *Avances en Educación y Humanidades*, *1*(1), 61-74. https://doi.org/10.21897/25394185.804

González-Gil F., Martín-Pastor, E., Flores Robaina, N., Jenaro Río, C., Poy Castro, R., & Gómez-Vela, M. (2013). Inclusion and school coexistence: analysis of teachers training, *European Journal of Investigation in Health, Psychology and Education*, *3*(2), 125–135. https://doi.org/10.1989/ejihpe.v3i2.30

Harvey, L. (2015). Beyond member-checking: A dialogic approach to the research interview. *International Journal of Research & Method in Education*, *38*, 23-38. https://doi.org/10.1080/1743727X.2014.914487

Hernández, M. J., & Carrasco, V. (2012). Percepciones de los estudiantes del máster en formación del profesorado de educación secundaria: fortalezas y debilidades del nuevo modelo formativo. *Enseñanza & Teaching*, *30*(2), 127-152.

Holstein, J. A., & Gubrium, J. F. (1995). *The Active Interview*. Sage.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal*, *72*(3), 283-294. https://doi.org/10.1111/j.1540-4781.1988.tb04190.x

Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. https://doi.org/10.1177/001316446002000116

McKelvie, S. J. (1978). Graphic rating scales: How many categories? *British Journal of Psychology*, *69*, 185-202. https://doi.org/10.1111/j.2044-8295.1978.tb01647.x

Medina, L., & Pérez, C. (2017). Adaptaciones metodológicas en el aula de lengua extranjera para el alumnado con discapacidad intelectual leve: estudio de caso. *Porta Linguarum*, (*2*), 267-282.

Muskat, M., Blackman, D. A., & Muskat, B. (2012). Mixed methods: Combining expert interviews, cross-impact analysis and scenario development. *The Electronic Journal of Business Research Methods*, *10*(1), 09-21. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2269508.

Navarro-Mateu, D., Franco-Ochoa, J., Valero-Moreno, S., & Prado-Gascó, V. (2020). Attitudes, sentiments, and concerns about inclusive education of teachers and teaching students in Spain. *Frontiers in Psychology*, *11*, 1-11. https://doi.org/10.3389/fpsyg.2020.00521

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage.

Reeves, J. (2004). 'Like everybody else': Equalizing educational opportunity for English language learners. *TESOL Quarterly*, *38*, 43-66. https://doi.org/10.2307/3588258

Reeves, J. R. (2006). Secondary teacher attitudes toward including English-language learners in mainstream classrooms. *The Journal of Educational Research*, *99*(3), 131-143. https://doi.org/10.3200/JOER.99.3.131-143

Reoyo, N., Carbonero, M. A., Freitas, A., & Valdivieso, J. A. (2012). La percepción de los futuros profesores sobre los docentes de educación secundaria. *International Journal of Developmental and Educational Psychology*, *2*(1), 389-395.

Rodríguez Izquierdo, R. (2005). Study on the understandings of teaching students about cultural diversity. *Educar*, *36*, 49. https://doi.org/10.5565/rev/educar.196

Roters, B., Nold, G., Haudeck, H., Keßler, J.-U., & Stancel-Piątak, A. (2011). Professionelles Wissen von Studierenden des Lehramts Englisch. In S. Blomeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert, & H. Willenberg (Eds.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen: Erste Ergebnisse aus TEDS-LT* (pp. 77-100). Waxmann.

Santos, J., & César, M. (2010). Atitudes e preocupacoes de professores e outros agentes educativos face à inclusao. *Interações*, *184*, 156-184. https://doi.org/10.25755/int.419

Savic, V., & Prosic-Santovac, D. (2017). English language teachers' attitudes towards inclusive education. *Inovacije u nastavi*, *30*(2), 141-157. https://doi.org/10.5937/inovacije1702141S

Schwab, S., Sharma, U., & Hoffmann, L. (2019). How inclusive are the teaching practices of my German, Maths and English teachers?—Psychometric properties of a newly developed scale to assess personalisation and differentiation in teaching practices. *International Journal of Inclusive Education*, *26*, 61-76. https://doi.org/10.1080/13603116.2019.1629121

Schwab, S., Tretter, T., & Gebhardt, M. (2013). Entwicklung und überprüfung eines fallbasierten instruments zur messung der einstellung zur schulischen integration. Wie denken Studierende, Berufstätige und Schüler/innen über schulische Integration von Kindern mit sonderpädagogischem Förderbedarf?. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, *83*(1), 20–32. https://doi.org/10.2378/vhn2014.art02d

Sharma, U., Forlin, C., Loreman, T., & Earle, C. (2006). Pre-service teachers' attitudes, concerns and sentiments about inclusive education: An international comparison of novice pre-service teachers. *International Journal of Special Education*, *21*(2), 80-93.

Taylor, R. W., & Ringlaben, R. P. (2012). Impacting pre-service teachers' attitudes toward inclusion. *Higher Education Studies*, *2*(3), 16-23. https://doi.org/10.5539/hes.v2n3p16

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, *19*, 349-357.

UNESCO. (1990). *World declaration on education for all*. UNESCO.

UNESCO. (1994). *The Salamanca Statement*. UNESCO.

UNESCO. (2015). *Education 2030 Incheon declaration: Towards inclusive and equitable quality education and lifelong learning for all*. UNESCO.

Vélez Latorre L. (2013). La educación inclusiva en docentes en formación: su evaluación a partir de la teoría de facetas. *Folios*, *37*, 95–113. https://doi.org/10.17227/01234870.37folios95.113