L-Università ta' Malta
Centre for Resilience & Socio-Emotional Health

International Journal of Emotional Education
ISSN: 2073 7629

Volume 16, Number 1, April 2024 pp 70 - 87

www.um.edu.mt/ijee

Assessing teachers' social and emotional competence: The validation of SECTRS in Italy, Latvia, and Portugal

Ilaria Grazzani[a] [1], Baiba Martinsone[b], Celeste Simoes[c], Valeria Cavioni[d], Elisabetta Conte[a], Veronica Ornaghi[a] and Alessandro Pepe[a]

[a] *University of Milano-Bicocca, Department of Human Sciences for Education 'Riccardo Massa', Italy*

[b] *University of Latvia, Department of Psychology, Latvia*

[c] *Universidade de Lisboa, Faculdade de Motricidade Humana, Portugal*

[d] *University of Foggia, Department of Social Sciences, Italy*

Despite growing interest in evaluating the social and emotional learning of students and the development of standardized assessment tools for this purpose, there remains a dearth of validated instruments for evaluating teachers' social and emotional skills. We set out to address this knowledge gap by investigating the psychometric characteristics of the Social-Emotional Competence Teacher Rating Scale (SECTRS) questionnaire, an instrument originally developed and validated in North America. Participants were 572 teachers from Italy (n=324), Latvia (n=139), and Portugal (n=109) ranging from 40 to 59 years of age. We performed confirmatory factor analysis procedures to validate a four-factor model measuring Teacher-student relationships, Emotion regulation, Social awareness, and Interpersonal relationships. The data supported this factorial structure, with a total of 14 items included in the final model. The outcomes of a multi-group comparison indicated that the model exhibited partial invariance, up to and including metric invariance, across the three cultural settings. We discuss these findings in relation to selecting appropriate instruments for assessing the social and emotional skills of teachers in different cultural contexts.

Keywords: SECTRS, SEL, social and emotional competence, assessment tool, teachers.

First submission 21st October 2023; Accepted for publication 1st March 2024.

1 Corresponding author. Email address: ilaria.grazzani@unimib.it

https://doi.org/10.56300/QIAN8168

ISSN 2073 7629

Introduction

In recent decades, there has been a growing focus on fostering and assessing the social and emotional competence (SEC) of children and adolescents, as well as their social and emotional learning (SEL). Studies on SEC have primarily been conducted within the field of developmental psychology. A recent work by Denham (2023) offers an excellent overview of the findings obtained with toddlers and preschoolers in almost 40 years of research. Denham focused on the three main components of emotional competence, namely emotional expressiveness, emotion knowledge, and emotion regulation, and how these abilities develop and may be evaluated in children. Denham further emphasized the complex connection between emotional competence and social and school success, pointing out that young children who understand and regulate their emotions are better at developing positive and supportive relationships with peers and educators and tend to be higher achievers (e.g., Cavioni et al., 2020a; Leerkes et al., 2008).

Based on these types of findings, the past two decades have seen the implementation in American schools of universal, comprehensive programs aimed at enhancing students' social and emotional learning. According to the CASEL (Collaborative on Academic Social and Emotional Learning) model, SEL comprises five competence, namely self-awareness, self-management, social awareness, relationship skills, and responsible decision making (Mahoney et al., 2021). Recent reviews and meta-analyses (Blewitt et al., 2021; Cefai et al., 2022) suggested that evidence-based SEL programs have a significant impact on each of these five competence and, in general, on students' mental health (e.g., Cavioni et al., 2023b; Taylor et al., 2017). This impact has been assessed using multiple validated measures such as the SSIS SEL Brief Scales (Anthony et al., 2020; Anthony et al., 2023; Anthony et al., 2022a; 2022b; Cavioni et al., 2023a; Elliott et al., 2020), the STQR (Student-Teacher Relationship Questionnaire, Murray & Greenberg, 2001), and the STRS (Student-Teacher Relationship Scale, Pianta, 2001).

As highlighted by various scholars (Cavioni et al., 2023b; Conte et al., 2023; Denham, 2023), teachers play a critical part in students' SEL outcomes because they are the primary implementers of SEL curricula (Taylor et al., 2017; Martinsone et al., 2022a). Furthermore, the teaching profession is considered one of the most emotionally demanding occupations. Thus, the preconditions for teachers to effectively implement social and emotional programs revolve around their own emotional skills and their well-being at school more generally. For example, when delivering interventions targeted at fostering and enhancing students' abilities in areas such as emotional self-awareness or emotion regulation strategies, teachers must either already possess adequate social and emotional competence of their own or undergo ad hoc training (Ornaghi et al., 2014). If this need is overlooked, teachers' levels of stress increase and students' academic achievement and behavior both suffer (Schonert-Reich, 2017).

Despite the increasing number of suitable instruments for assessing students' SEL, there is currently a lack of validated measures for evaluating the social and emotional competence of the teachers charged with implementing SEL programs in schools. Indeed, as recently shown by Lozano-Pena et al. (2021), teachers' social and emotional competence have almost exclusively been evaluated with respect to the construct of

emotional intelligence (see for example, Mayer & Salovey, 1997). Only recently has the CASEL framework prompted the assessment of teachers as well as students, and this primarily in the United States. One of the few existing instruments that explicitly measures teachers' social and emotional skills is the Yoder (2014) questionnaire. This tool assesses the five dimensions in the CASEL framework but is not available free of charge. More recently, Elliott and a group of colleagues (e.g., Anthony et al., 2020) have begun to develop a comprehensive battery of scales for future measurements of both student and teacher SEL skills.

In the European context, the CASEL framework has only begun to take off in recent years, in terms of the development of validated programs and assessment tools. It follows that the current lack of measures for evaluating teachers' SEL – that is their ability to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions – is particularly marked in Europe. Indeed, although SEL programs for teachers are growing in number and rapidly becoming a key focus within European education systems (e.g., Agliati et al., 2020; Cavioni et al., 2023b; Madalinska-Michalak, 2015), to the best of our knowledge, there are currently no validated tools for assessing European teachers' social and emotional competence.

In light of this background, we present a study conducted within a large, evidence-based European research project, the Promehs, which was designed to promote school mental health with a special focus on social and emotional learning (for background on the Promehs project, see Cefai et al., 2022). The aim of our research was to validate a questionnaire developed in the United States for assessing teachers' social and emotional competence, as part of a broader set of validation studies of American tools (see, for example, Cavioni et al., 2023a) administered under the Promehs project. Specifically, we report on the validation of the SECTRS questionnaire (Tom, 2012) in Italy, Latvia, and Portugal. These three countries, which were part of the European network set up under the Promehs project, were those able to recruit a large enough sample size to satisfactorily power the validation analyses.

In the next sections, we describe the SECTRS instrument, as well as our sample of participants, data collection procedures, and data analysis strategy. Finally, we report the outcomes of the study, discussing both their implications and limitations.

The SECTRS

The Social-Emotional Competence Teacher Rating Scale (SECTRS; Tom, 2012) was developed exclusively for the assessment of teachers' social and emotional competence. The author devised the instrument as part of her doctoral research and built the items of the scale with the input of an expert panel. She carried out the study with 302 teachers from primary (grade K-2), elementary (grade 3-5), middle (grade 6-8) and high school (grade 9-12). Following an iterative content validation process, the scale's factor structure, reliability and validity were explored in light of the demographic characteristics of the sample. The final version of the original SECTRS instrument comprised 52 items measuring four factors in teachers' social and emotional competence, namely Teacher-student relationships, Emotion regulation, Social awareness, and Interpersonal relationships. The Teacher-student relationships items tap into the quality of the interactions between teacher and students

(e.g., *I am aware of how all of my students are feeling*); the Emotion regulation items are focused on teachers' ability to manage their own emotions in challenging situations in the classroom (e.g., *I nearly always stay calm when a student upsets me*); the Social awareness subscale concerns teachers' ability to reflect on how their own decisions and behaviors impact students, as well their acceptance of diversity (e.g., *I work well with students from diverse backgrounds*); finally, the Interpersonal relationships sub-scale evaluates teachers' relationships with other school staff and the parents of their students (i.e., *I build positive relationships with my students' families)*. Each item is rated on a 6-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). Based on her initial findings, Tom (2012) concluded that SECTRS was sufficiently psychometrically robust; nevertheless, she recommended further refinement and validation of the instrument before using it in other research and clinical practice settings.

To the best of our knowledge, the SECTS has only been adapted and validated for Mexican samples (Moncalvo Curiel, et al., 2018, 2019). This study, which involved 222 teachers, assessed the scale's reliability, construct validity, and factorial structure. Construct validity was satisfactory and each of the subscales attained reliability indices of $\alpha \geq .70$, which was comparable to the level of reliability observed for the original English version.

The Present Study

The aim of the present study was to validate the structural integrity of the measurement model for the Social-Emotional Competence Teacher Rating Scales (SECTRS). To this end, we drew on confirmatory factor analysis methods to assess various models, using data collected with the questionnaire in three different cultural contexts. Specifically, we examined four different models of the SECTRS: a one-dimensional model (M1), a four-dimensional model with independent scales and restrictions on the covariance of item-level errors (M2), a four-dimensional model with dependent scales and restrictions on the covariance of item-level errors (M3), and a four-dimensional model with dependent scales and correlated uniqueness of items (M3a).

Given that the dimensions of SEC evaluated by the questionnaire (i.e., teacher-student relationships, emotion regulation, social awareness, and interpersonal relationships) are interconnected, it is reasonable to expect correlations among them, potentially strong ones, as predicted by existing empirical evidence (e.g., Barthel et al., 2018; Niven, 2017; Pallini et al., 2019; Sutton et al., 2009). Furthermore, because this study adopted a multi-contextual approach, we expected to find instances of covariance in item-level errors. This would be justified by the fact that correlated item-level errors can capture local variations in data patterns (Hermida, 2015).

To enhance the generalizability of our findings, we also performed multiple-group confirmatory factor analysis (MGCFA), as well as a cross-validation procedure, with a view to testing measurement invariance across three national datasets (i.e., Italy, Latvia, and Portugal). The hypothesis of measurement invariance was considered valid if configural invariance (a weak form of factorial invariance, Horn & McArdle, 1992), metric invariance, residual variance invariance, and full construct invariance were supported (for more details, see Cheung & Rensvold, 2002). Given the need to realistically compare latent constructs across different groups,

it is essential that observed scores exhibit consistent patterns across all groups. Because psychological studies often involve comparing constructs across different cultural cohorts (Vijver & Leung, 2000), measurement invariance is a key consideration when mapping the psychological characteristics of in-service teachers.

Method

Participants

The sample comprised 572 teachers from three countries, namely Italy (n=324), Latvia (n=139) and Portugal (n=109). Participants were enlisted to the study at 21 Italian schools, 15 Latvian schools, and 29 Portuguese schools, either private or public, located in urban and rural areas, via direct contacts or email, telephone, and social media channels, from kindergarten to high schools. The decision to take part in the study was based on the participants' own interest in taking part in a European project on promoting school mental health (Cefai et al., 2022). In all three countries, more than 90% of participants were female teachers.

The inclusion criteria for the study were: (1) being a professionally trained in-service teacher and (2) agreeing to the terms and conditions of participation in the research. No particular exclusion criteria were applied. We gathered a convenience sample using a non-probability sampling technique (Emerson, 2015). The data was collected anonymously, and all participants were made aware of the study's aims and methods. The Ethics Board of the University of Milano-Bicocca (Protocol number: 0044281/20 obtained on the 21st/7/July 2020) approved the research, which was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki (World Medical Association, 2000) and the American Psychological Association code of conduct (Knapp & Vande Creek, 2004).

Measures and procedures

The teachers completed either a pencil and paper version or an online version of the SECTRS (Social-Emotional Competence Teacher Rating Scale; Tom, 2012). The instrument was introduced as follows: "The statements below describe your thoughts, feelings, and actions in the classroom and in situations at your school. For each item, please indicate the extent to which you agree or disagree with the statement provided. There are no right or wrong answers, so please be as honest as possible". The questionnaire consisted of 52 items to be rated on a 6-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree) and divided into the following sub-scales: Teacher-student relationships, Emotion regulation, Social awareness, and Interpersonal relationships.

Statistical Insights: Data Analysis

In validating a measurement model for the Social-Emotional Competence Teacher Rating Scales (SECTRS) based on data from the three cohorts of European teachers, we followed multivariate Confirmatory Factor Analysis (CFA) procedures as in Cavioni et al. (2020b). This type of analysis involves comparing the hypothetical covariance matrix reproduced by the model (Σ) with the effectively observed empirical data (S)

(Kline, 2013): the degree of overlap between the two matrices reflects the extent to which the model offers a suitable fit for the data.

First, we examined the data to assess whether the scores for each of the items were normally distributed. None of the score sets had kurtosis or skewness values that exceeded the recommended limits of 2 to 4 for the former or -1 to 1 for the latter (Blanca et al., 2013). We next assessed the data for multivariate outliers using Mahalanobis distances (Ghorbani, 2019). As a result, five cases were removed from the Italian sample, but none from the Portuguese or Latvian samples; the retention rate was equal to 98.5% for the Italian cohort and 100% for the others. The Maximum Likelihood method was used to estimate the SEM model (Ripplinger & Sullivan, 2008).

To estimate model fit, we calculated five goodness-of-fit indices, namely the Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Normed Fit Index (NFI), Tucker-Lewis Index (TLI, TLI >.95), and Comparative Fit Index. The model was considered to fit the data if the RMSEA was less than .07 (Kenny et al., 2015), the SRMR less than.05 (Ximénez et al., 2022), and the NFI greater than .95 (Shi & Maydeu-Olivares, 2020). We estimated confidence limits with a set of 200 random samples in keeping with current SEM practices (Thakkar, 2020). Given that we planned to compare different measurement models for the SECTRS, we also adopted Akaike's information criterion (AIC), with lower values indicating a better fit (Bozgodan, 1987).

Finally, we tested structural invariance across culture-based groups using multi-group CFA (MGCFA; Byrne, 1998). The hypothesis of group invariance was to be accepted if configural invariance (i.e., the underlying relationships between variables are stable across groups), metric invariance (i.e., the model parameters, such as regression coefficients, bear the same meaning across groups), and scalar invariance (i.e., the unit of measurement is the same across groups) were supported (see Van De Schoot et al., 2015). Equivalence of the model across groups was rejected if the difference between the target model and the nested models was statistically significant. We set the cutoff criteria for rejecting invariance at $\Delta > 0.01$ for both ΔRMSEA and ΔSRMR (Chen, 2007) and a chi-square difference ($\Delta\chi 2$) that was statistically significant at the $p < .01$ level (Milfont & Fischer, 2010). The different types of invariances are hierarchically ordered, meaning that the MGCFA procedure ends at the lowest level of invariance that fails to be satisfied (for further details, see Cheung & Rensvold, 2002).

Results

Confirmatory Factor Analysis: setting up a baseline model

Prior to examining measurement invariance across groups, we used a confirmatory technique to identify a baseline model that offered a good fit for the data. We used the largest sample as our reference group to maximize the statistical robustness of our analysis. We assessed four distinct models for the SECTRS with the aim of identifying that best aligned with the observed data. The outcomes are presented in Table I.

Table I

Outcomes of Four Distinct Models for the SECTRS

χ^2 (df)	NFI	NNFI	CFI	RMSEA	AIC	Decision
1038.4 (275)	.673	.710	.734	.143	1188.4	Reject
1188.7 (275)	.626	.653	.682	.101	1338.7	Reject
656.8 (269)	.743	.850	.865	.067	818.8	Reject
581.1 (260)	.817	.876	.890	.061	749.2	Reject

The initial confirmatory factor analyses (CFA) indicated that data failed to fit any of the models assessed (M1-M3a). Upon further examination of the structural values, it was observed that a number of the standardized factor loadings ($\lambda i,j$) fell below the threshold of .05 recommended by Kim and Yoon (2011). Consequently, only the items that met the cut-off criterion were retained and the hypothesized models were retested using the reduced set of items. The outcomes are fully reported in Table II.

Table II

Results of Reduced (14 items) Structural Model on the Reference Group (i.e., Italian sample)

Model	χ^2 (df)	NFI	NNFI	CFI	RMSEA	AIC	Decision
M1. Unidimensional	490.8 (77)	.746	.735	.775	.129	574.8	Reject
M2. Four-dimensional, independent factors	629.5 (77)	.675	.646	.700	.149	713.5	Reject
M3. Four-dimensional, dependent factors, uncorrelated errors	188.6 (71)	.903	.918	.936	.072	284.6	Reject
M3a. Four-dimensional, dependent factors, correlated errors	155.1 (68)	.920	.953	.953	.063	257.1	Accept

The confirmatory factor analysis (CFA) results suggested that M1 did not fit the data and should be rejected. M2 modeled the notion that the various dimensions of social and emotional competence in teachers are unrelated to each other. This model also offered a suboptimal fit for the data and was rejected. M3 was similar in structure to M2 but reflected the interdependence among the four dimensions of Teacher-student relationships, Emotion regulation, Social awareness, and Interpersonal relationships. With respect to M2, the goodness-of-fit indicators for M3 were somewhat better, but the values for CFI and NNFI did not meet the predetermined cut-off criteria. The final model, M3a offered the best level of fit and obtained the lowest AIC value of all the models. The data was fitted to the null model, resulting in a $\chi 2(18)$ value of 155.1, which was statistically significant ($p < .01$). The model also had an RMSEA value of .063, a p close value of .157, an NFI value of .920, a NNFI value of .953, a CFI value of .953.This prompted full acceptance of M3a as a baseline model for further analyses (Table III). Indeed, overall, our findings offered empirical support for the proposed SECTRS measurement model, which consists of four distinct dimensions that are closely interconnected and

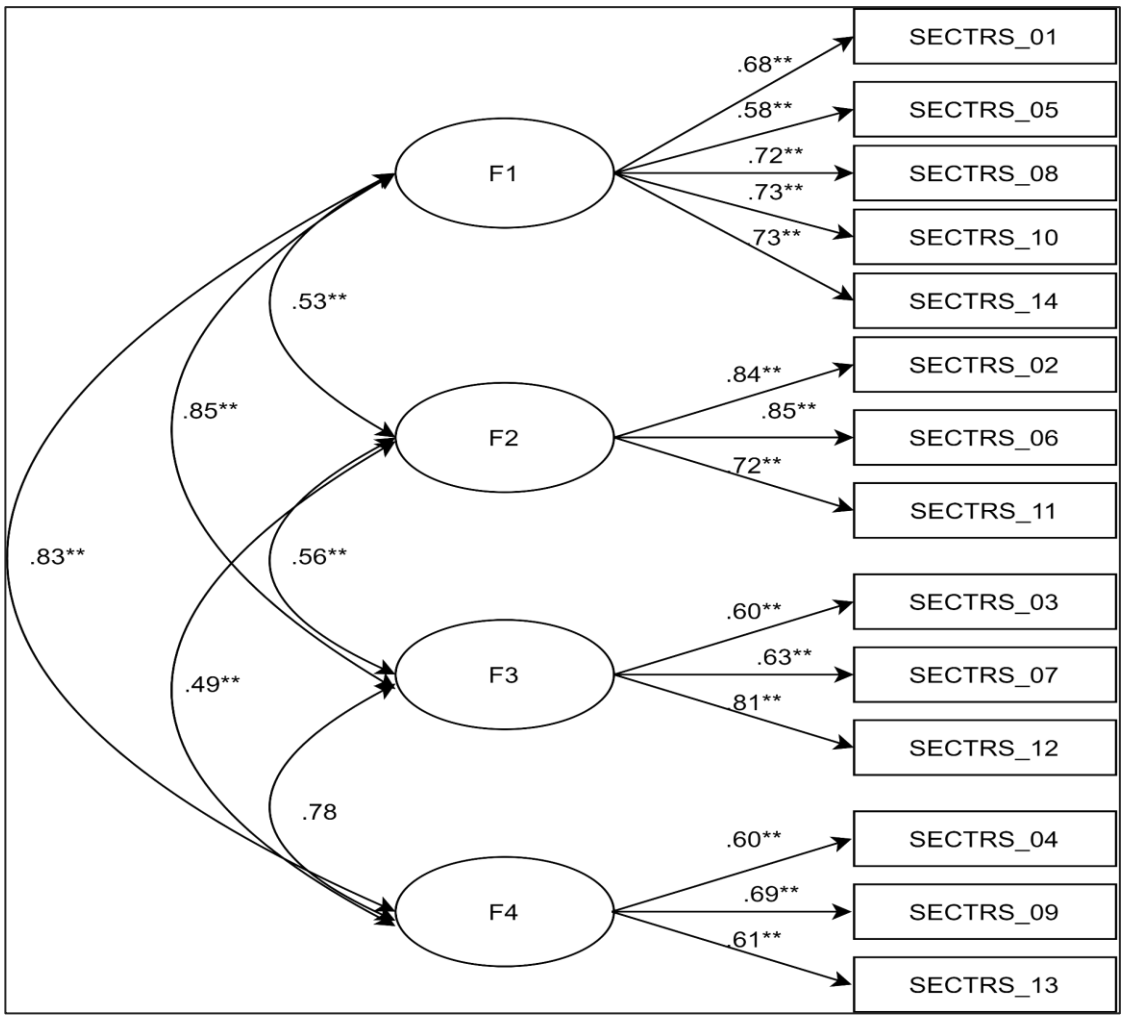
14 items. All the indicators exhibited values that were above the acceptable threshold. The factor loadings are presented in Figure 1.

Measurement Invariance of SECTRS: Multigroup Comparison

Next, we performed MGCFA to assess the measurement invariance of SECTRS across the three datasets (e.g., Italian, Latvian and Portuguese). In essence, MGCFA enables the assessment of the degree to which individuals belonging to distinct cohorts attribute equivalent meaning to items of a scale (Gouveia, et al., 2009). Following the standard protocol, we initially evaluated configural invariance and subsequently examined additional, more complex forms of equivalence.

Figure 1

Structural equation model for the factor structure of the SECTRS.



Standardized loadings were reported. F1=Teacher-student Relationship (α = .817), F2= Emotion Regulation (α = .828), F3= Social Awareness (α = .699), F4= Interpersonal Relationship (α = .625). **p < .01

Table III

Results of Structural Model Comparison of SECTRS between Italy, Latvia, and Portugal.

Model	χ^2 (df)	CFI	RMSEA	[90% CI]	Model comparison	Δχ2 (Δdf)	ΔCFI	ΔRMSEA	Decision
M1. Configural invariance	364.2 (204)	.947	.037	0.031 - 0.044	-	-	-	-	Accept
M2. Metric invariance*	388.5 (220)	.944	.038	0.031 - 0.043	Ma	24.3 (16)	.003	.001	Accept
M3. Scalar invariance	661.2 (252)	.865	.054	0.044 - 0.059	Mb	272.7 (32)	.121	.016	Reject
M3a. Full invariance	697.6 (272)	.859	.053	0.048 - 0.058	Mc	36.4 (20)	.006	.001	Reject
Note. Metric invariance was reached by using items 1 and 7 as unconstrained indicators.									

First, we tested for configural invariance across groups. The model was found to provide a satisfactory fit for the data, supporting the hypothesis of configural invariance. Next, metric invariance (whereby factor loadings for items and constructs are constrained to be equal across groups) was tested, with the outcomes suggesting that some of the items assessed the same underlying construct across groups. Differential analysis of structural parameters suggested that, in the Portuguese sample, Item 01 "I have a close relationship with my students" and Item 07 "I problem-solve with students when there is a problem or argument" should be set as free to vary. The outcomes of the MGCFA essentially confirmed metric invariance, implying that it is appropriate to use the measure to test for differences between groups. Third, we assessed scalar invariance. In this case, the model did not provide a good fit for the data. Finally, we tested for full invariance (whereby items display similar uniqueness in relation to the underlying constructs across the different groups), again obtaining model fit indicators that fell outside the acceptable ranges, resulting in a $\chi2(18)$ value of 697.6, which was statistically significant ($p < .01$). The model also had an RMSEA value of .053 and all the comparisons with the previous model did not support the acceptance of the invariance level. This implied that the SECTRS (referred to from now on as SECTRS-14; see Appendix 1) measurement model was not fully equivalent across different cohorts of teachers.

Discussion

Methods for evaluating teacher's social and emotional competence (SEC) are required to investigate the influence of this variable on the implementation of SEL programs and on academic outcomes, among other research and applied purposes. Hence, the aim of this study was to validate the SECTRS, an American tool (Tom, 2012) designed to measure teachers' social and emotional competence. This involved testing different factor structures in search of a statistically sound factor solution, for samples from Italy, Latvia, and Portugal. Given that social and emotional competence include culturally-situated skills (e.g., Cavioni et al., 2023a; Martinsone et al., 2022b), it is crucial to develop questionnaires that are attuned to the social and emotional competence of teachers within specific cultural settings.

The results of this study offered empirical support for a 14-item version of SECTRS for use across three European countries - Italy, Latvia, and Portugal - among those involved in the Promehs Project network. This shorter version of the SECTRS questionnaire is simpler and more user-friendly than either the original instrument or the previously published Mexican adaptation (Moncalvo Curiel et al., 2019). As in the original scale, the SECTRS-14 assesses four interconnected factors, namely teacher-student interaction, emotion management, social awareness, and interpersonal relationships. Although these factors are highly correlated, they are distinct and different from each other. Thus, a 14-item SECTRS would allow teachers' social and emotional competence to be assessed via a single brief tool that taps into a range of key components of SEC, from emotion regulation to social awareness.

An agile and, at the same time, robustly validated tool could be especially valuable for assessing the effectiveness of training and intervention programs from a longitudinal perspective. Recent years have seen an increase in intervention studies aimed at evaluating the efficacy of programs targeting not only students but also teachers' school mental health and wellbeing (e.g., Cavioni et al., 2023b for a recent review). Evaluating such outcomes, as researchers have often highlighted, has been problematic given the lack of validated and easy-to-complete assessment tools. Hence, the SECTRS-14 has the potential to serve as a valuable instrument for researchers and policy makers interested in evaluating longitudinal changes in teacher SEC as a function of specific variables such as the implementation of training courses aimed at enhancing teachers' social and emotional abilities.

We found the SECTRS measurement model displayed partial invariance across cultural cohorts of teachers from Italy, Latvia, and Portugal, although there were significant unaccounted-for differences. There are various possible explanations for this failure to attain complete invariance. First, it could be due to cultural differences in how teachers understand and display their social and emotional skills. Different cultural backgrounds may emphasize in different ways the role of specific social and emotional abilities, resulting in differences in how individual teachers fill out the questionnaire. Second, differences across countries in education systems and practices (e.g., teaching methods, institutional support etc.) may also impact teachers' social and emotional abilities. Third, linguistic discrepancies between the original questionnaire, which was devised and validated in North America, and the European versions developed for this study may have led to incomplete equivalency. Indeed, translating and adapting assessment tools for use across various languages and cultural contexts can add or remove subtleties and details that can affect the equivalence of the final instrument.

In brief, while the SECTRS questionnaire displayed potential in evaluating teachers' social and emotional competencies across different cultural settings, the partial invariance finding underscores the importance of acknowledging and addressing cultural and contextual differences when using such assessment tools in global research and practice. Future research should explore these differences more thoroughly and improve the measurement methodology to increase its cross-cultural validity and usefulness.

Conclusion and Limitations

The aim of this study was to address a significant gap in our understanding of how to assess the social and emotional skills of teachers. By focusing on the psychometric properties of the Social-Emotional Competence Teacher Rating Scale (SECTRS) questionnaire across different European educational settings, and involving teachers from Italy, Latvia, and Portugal, we provided insights into the effectiveness of this tool. Confirmatory factor analysis enabled us to validate a four-factor structure comprising Teacher-student relationships, Emotion regulation, Social awareness, and Interpersonal relationships, as measured via a set of 14 items. Importantly, we found that while there were some differences among the different cultural settings, the questionnaire displayed partial invariance in its reliability and validity across diverse contexts. These results confirm the value of deploying robust assessment instruments such as the SECTRS to assess teachers' social and emotional competence with a view to enhancing educational practices across settings.

Nevertheless, the study displays a number of limitations. First, reducing the SECTRS questionnaire to 14 items may raise concerns surrounding changes to the instrument's features and imply the need for external validation. It should be noted that the reduction was data-driven and also of potential value in terms of making the instrument easier to administer in future studies. At the same time, decreasing the item count could call the questionnaire's comprehensiveness and reliability into question. Changing the initial design of the instrument could impact its continued accuracy and applicability. Hence, it is crucial to recognize the possible consequences of this decrease and to recommend external validation tests with a view to ensuring the ongoing psychometric integrity of the redesigned tool. Indeed, given the cross-sectional nature of the current study, future longitudinal investigations are required to confirm the stability of the outcomes, preferably using an even larger sample.

A second limitation is the difference in the sizes of the three cultural samples. Italy recruited more participants to the investigation than did Latvia and Portugal, and this could generate bias in the findings. Given that cultural differences, educational methodologies, and social expectations vary among countries, the overrepresentation of a particular cultural group could disproportionately influence the outcomes of the investigation and limit the applicability of the results to different cultural settings. Hence, it is essential to recognize this constraint and evaluate its possible influence on the present findings. Future studies should aim for a more equitable sample distribution to enhance the robustness and accuracy of the findings across cultural settings.

Third, we also acknowledge sample size as a limitation in relation to some of the specific analyses that we conducted. Adequate sample size is crucial in study design, especially when performing intricate analyses such as confirmatory factor analysis (CFA) or when comparing multiple groups from different cultural contexts. Furthermore, although attempts were made to recruit individuals from a range of backgrounds within Italy, Latvia, and Portugal, respectively, the sample size from each of the individual countries may have been too small to fully represent the internal diversity of each cultural group. It is crucial to recognize the possible influence of this constraint on the statistical power and reliability of our analyses. A larger and more diversified sample size would have provided a more thorough investigation of cultural variations and a deeper insight into

the psychometric characteristics of the SECTRS questionnaire. Future studies should be focused on recruiting larger and more diverse samples to enhance the validity and generalizability of the results. Researchers should use additional methods or perform further investigations to examine and confirm the current pattern of findings with larger and more diverse groups of participants. Future studies can offer more reliable findings by overcoming the sample size limitation incurred here, thus expanding our understanding of teachers' social and emotional abilities across different cultural scenarios.

A final limitation concerns the characteristics of the participating school systems. We should be aware that there are contextual variables linked to specific school systems that can impact the interpretation of findings in comparative research designs. In other words, the teachers' scores on SECTRS-14 could vary as a function of numerous factors, including the fact that the student population of a particular school might be particularly problematic (low SES background, widespread psychological distress in the community, a high number of students with disabilities and/or special learning needs, a lack of material and financial resources, etc.). Therefore, teachers' scores on the social and emotional competence scale may differ significantly when collected in schools and socio-cultural contexts that subject teachers to intense emotional stress and put them at risk of burnout. Hence, we should recognize the potential for regional bias in our sample, as well as the fact that some groups of teachers may be underrepresented, which could weaken the generalizability of the findings.

Even with these limitations in mind, the present study represents a promising contribution to a line of inquiry aimed at validating self-report tools for assessing teachers' social and emotional abilities across different European countries.

Acknowledgements

This work was supported by Erasmus+ KA3 (PROMEHS Project N°0.606689-EPP-1-2018-2-IT-EPPKA3-PI-POLICY). The Project was coordinated by the University of Milano-Bicocca. We are very grateful to the researchers from the three countries who have been involved in the Project. They have contributed significantly to the translation of the questionnaire and to the data collection.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial and financial relationships that could be construed as a potential conflict of interest.

References

Agliati, A., Barriga, P., Cifuentes, P., Baena, I., Berzanskyté, J., Cavioni, V., Conte, E., Cuadrado Méndez, F., Ferreira, M., Gnezda Fajfar, A., Gómez Baya, D., Grazzani, I., Lazdina, S., López Cobo, I., Menor Campos, E., Martínez Jiménez, E., Martinsone, B., Muñiz Velázquez, J. A., Ornaghi, V., & Valverde Jiménez, B. (2020). *Learning to be: Toolkit for assessing social and emotional skills at school.* (CC BY-NC 4.0). Erasmus+ Programme of the European Union.

Anthony, C. J., Brann, K. L., Elliott, S. N., & Garis, E. J. (2022a). Examining the structural validity of the SSIS SEL brief scales-Teacher and student forms. *Psychology in the Schools, 59*(2), 260-280. https://doi.org/10.1002/pits.22607

Anthony, C. J., Elliott, S. N., DiPerna, J. C., & Lei, P.-W. (2020). The SSIS SEL brief scales–student form: Initial development and validation. *School Psychology Forum, 35*(4), 277-283. https://psycnet.apa.org/doi/10.1037/spq0000390

Anthony, C. J., Elliott, S. N., Yost, M., Lei, P. W., DiPerna, J. C., Cefai, C., Camilleri, L., Bartolo, P. A., Grazzani, I., Ornaghi, V., Cavioni, V., Conte, E., Tatalović Vorkapić, S., Poulou, M., Martinsone, B., Simões, C., & Colomeischi, A. A. (2022b). Multi-informant validity evidence for the SSIS SEL brief scales across six European countries. *Frontiers in Psychology, 13,* 928189. https://doi.org/10.3389/fpsyg.2022.928189

Anthony, C. J., Lei, P.-W., Elliott, S. N., DiPerna, J. C., Cefai, C., Bartolo, P. A., Camilleri, L., O'Riordan, M., Grazzani, I., Cavioni, V., Conte, E., Ornaghi, V., Tatalović Vorkapić, S., Poulou, M., Martinsone, B., Simões, C., & Colomeischi, A. A. (2023). Measurement invariance of children's SEL competencies: An examination of SSIS SEL brief scales with a multi-informant sample from six countries. *European Journal of Psychological Assessment.* https://doi.org/10.1027/1015-5759/a000753

Barthel, A. L., Hay, A., Doan, S. N., & Hofmann, S. G. (2018). Interpersonal emotion regulation: A review of social and developmental components. *Behaviour Change, 35*(4), 203-216. https://doi.org/10.1017/bec.2018.19

Blanca, M. J., Arnau, J., López-Montiel, D., Bono, R., & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology, 9*(2). https://doi.org/10.1027/1614-2241/a000057

Blewitt, C., O'Connor, A., Morris, H., May, T., Mousa, A., Bergmeier, H., Nolan, A., Jackson, K., Barrett, H., & Skouteris, H. (2021). A systematic review of targeted social and emotional learning interventions in early childhood education and care settings. *Early Child Development and Care, 191*(14), 2159-2187. https://doi.org/10.1080/03004430.2019.1702037

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*, 345-370. https://doi.org/10.1007/BF02294361

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Lawrence Erlbaum.

Cavioni, V., Conte, E., Grazzani, I., Ornaghi, V., Cefai, C., Anthony, C., Elliott, S. N., & Pepe, A. (2023a). Validation of Italian students' self-ratings on the SSIS SEL brief scales. *Frontiers in Psychology, 14*, 1229653. https://doi.org/10.3389/fpsyg.2023.1229653

Cavioni, V., Grazzani, I., Ornaghi, V., Agliati, A., Gandellini, S., Cefai, C., Camilleri, L., Bartolo, P., Tatalović Vorkapić, S., Golob, L., Poulou, M., Martinsone, B., Supe, I., Simões, C., Lebre, P., Colomeischi, A., Rusu, P., Acostoaie, L., Vintur, T., & Conte, E. (2023b). A multi-component

curriculum to promote teachers' mental health: Findings from the PROMEHS program. *International Journal of Emotional Education, 15*(1), 34-52. https://doi.org/10.56300/KFNZ2526.

Cavioni, V., Grazzani, I., Ornaghi, V., Pepe, A., & Pons, F. (2020b). Assessing the factor structure and measurement invariance of the Test of Emotion Comprehension (TEC): A large cross-sectional study with children aged 3-10 years. *Journal of Cognition and Development, 21*(3), 406-424. https://doi.org/10.1080/15248372.2020.1741365

Cavioni, V., Grazzani, V., Ornaghi, V., Agliati, A., & Pepe, A. (2020a). Adolescents' mental health at school: The mediating role of life satisfaction. *Frontiers in Psychology, 12,* 770628. https://doi: 10.3389/fpsyg.2021.720628

Cefai, C., Camilleri, L., Bartolo, P., Grazzani, I, Cavioni, V., Conte, E., Ornaghi, V., Agliati, A., Gandellini, S., Tatalović Vorkapić, S., Poulou, M., Martinsone, B., Stokenberga, I., Simões, C., Santos, M., & Colomeischi, A. (2022). The effectiveness of a school-based, universal mental health programme in six European countries. *Frontiers in Psychology. 13,* 925614. https://doi.org/10.3389/fpsyg.2022.925614.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Conte, E., Cavioni, V., Ornaghi, V., Agliati, A., Gandellini, S., Santos, M. F., Santos, A. C., Simoes, C., & Grazzani, I. (2023). Supporting preschoolers' mental health and academic learning through the PROMEHS Program: A training study. *Children, 10,* 1070. https://doi.org/10.3390/children10061070

Denham, S. A. (2023). *The development of emotional competence in young children*. The Guilford Press.

Elliott, S. N., Anthony, C. J., DiPerna, J. C., Lei, P. W., & Gresham, F. M. (2020). *SSIS SEL brief scales: User guide and technical manual*. SAIL CoLab.

Emerson, R. W. (2015). Convenience sampling, random sampling, and snowball sampling: How does sampling affect the validity of research? *Journal of Visual Impairment & Blindness, 109*(2), 164-168. https://doi.org/10.1177/0145482X15109002

Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis, Series: Mathematics and Informatics*, *34*(3), 583-595. https://doi.org/10.22190/FUMI1903583G

Gouveia, V. V., Milfont, T. L., Da Fonseca, P. N., & Coelho, J. A. P. D. M. (2009). Life satisfaction in Brazil: Testing the psychometric properties of the satisfaction with life scale (SWLS) in five Brazilian samples. *Social Indicators Research, 90,* 267-277. https://doi.org/10.1007/s11205-008-9257-0

Hermida, R. (2015). The problem of allowing correlated errors in structural equation modeling: Concerns and considerations. *Computational Methods in Social Sciences, 3*(1), 5-17.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144. https://doi.org/10.1080/03610739208253916

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486-507. https://doi.org/10.1177/0049124114543236

Kim, E. S., & Yoon, M. (2011). Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. *Structural Equation Modeling, 18*(2), 212-228. https://doi.org/10.1080/10705511.2011.557337

Kline R. B. (2013). *Exploratory and confirmatory factor analysis*. In Y. Petscher, C. Schatsschneider,. & D. L. Compton (Eds.), *Applied quantitative analysis in the social sciences* (pp. 171-220). Routledge.

Knapp, S., & VandeCreek, L. (2004). A principle-based analysis of the 2002 American Psychological Association ethics code. *Psychotherapy: Theory, Research, Practice, Training, 41*(3), 247-254. https://doi.org/10.1037/0033-3204.41.3.247

Leerkes, E. M., Paradise, M., O'Brien, M., Calkins, S. D., & Lange, G. (2008). Emotion and cognition processes in preschool children. *Merrill-Palmer Quarterly, 54*(1), 102-124. https://psycnet.apa.org/doi/10.1353/mpq.2008.0009

Lozano-Peña, G., Sáez-Delgrado, F., Lopez-Angulo, Y., & Mella-Norambuena, J. (2021). Teachers' social-emotional competence: History, concept, models, instruments and recommendations for educational quality. *Sustainability, 13,* 12142. https://doi.org/10.3390/su132112142

Madalinska-Michalak, J. (2015). Development of emotional competence for teaching. *Croatian Journal of Education, 17*(2), 71-97.

Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R.J., Niemi, K., Schlinger, M., Schlund, J., Shriver, T. P., VanAusdal, K., & Yoder, N. (2021). Systemic social and emotional learning: Promoting educational success for all preschool to high school students. *American Psychologist, 76*(7), 1128-1142. https://doi.org/10.1037/amp0000701

Martinsone, B., Stokenberga, I., & Grazzani, I. (2022a). Monitoring system of implementation of the Promoting Mental Health at Schools (PROMEHS) program. *Frontiers in Psychology, 13,* 1430001. https://doi.org/10.3389/fpsyg.2022.1043001

Martinsone, B., Stokenberga, I., Damberga, I., Supe, I., Simoes C., Lebre, P., … Camilleri, L. (2022b). Adolescent social emotional skills, resilience and behavioral problems during the COVID-19 pandemic: A longitudial study in three European countries. *Frontiers in Psychiatry, 13,* 942692. https://doi.org/10.3389/fpsyt.2022.942692

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), Emotional development and emotional intelligence: Educational implications (pp. 3-31). Basic Books.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111-130. https://doi.org/10.21500/20112084.857

Moncalvo Curiel, A., Camacho, M., & Corral-Frías, N. S. (2018). Adaptación al español del cuestionario competencia socioemocional escala de calificación docente**.** *Revista de Comunicación y Salud, 8*(1), 61-69 https://doi.org/10.35669/revistadecomunicacionysalud.2018.8%281%29.61-69

Moncalvo Curiel, A., Castillo, K. N. F., Camacho, M., & Corral-Frías, N. S. (2019). Validación de una adaptación al español de la Escala de Competencia Socioemocional para Docentes. *Psicumex, 9*(1), 18-36. https://doi.org/10.36793/psicumex.v9i1.182

Murray, C., & Greenberg, M. T. (2001). Relationships with teachers and bonds with school: Social emotional adjustment correlates for children with and without disabilities. *Psychology in the Schools, 38*(1), 25–41. https://doi.org/10.1002/1520-6807(200101)38:1%3C25::AID-PITS4%3E3.0.CO;2-C

Niven, K. (2017). The four key characteristics of interpersonal emotion regulation. *Current Opinion in Psychology, 17,* 89-93. https://doi.org/10.1016/j.copsyc.2017.06.015

Ornaghi, V., Brockmeier, J., & Grazzani, I. (2014). Enhancing social cognition by training children in emotion understanding: A primary school study. Journal of Experimental *Child Psychology, 119,* 26-39.

Pallini, S., Vecchio, G. M., Baiocco, R., Schneider, B. H., & Laghi, F. (2019). Student–teacher relationships and attention problems in school-aged children: The mediating role of emotion regulation. *School Mental Health, 11,* 309-320. https://doi.org/10.1007/s12310-018-9286-z

Pianta R. C. (2001). *Student-Teacher Relationship Scale: Professional manual.* Psychological Assessment Resources, Inc.

Ripplinger, J., & Sullivan, J. (2008). Does choice in model selection affect maximum likelihood analysis? *Systematic Biology, 57*(1), 76-85. https://doi.org/10.1080/10635150801898920

Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Future for Children, 27*(1), 137-155.

Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement, 80*(3), 421-445. https://doi.org/10.1177/0013164419885164

Sutton, R. E., Mudrey-Camino, R., & Knight, C. C. (2009). Teachers' emotion regulation and classroom management. *Theory into Practice, 48*(2), 130-137.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow up effects. *Child Development, 88*(4), 1156-1171. https://doi.org/10.1111/cdev.12864

Thakkar, J. J. (2020). *Structural equation modelling: Application for research and practice.* Springer.

Tom, K. M. (2012). Measurement of teachers' social-emotional competence: Development of the Social-emotional Competence Teacher Rating Scale. https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/12351/Tom_oregon_0171A_10250.pdf?sequence=1&isAllowed=y

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Measurement invariance. *Frontiers in Psychology, 6,* 1064. https://doi.org/10.3389/fpsyg.2015.01064

van de Vijver, F. J. R., & Leung, K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology, 31*(1), 33-51. https://doi.org/10.1177/0022022100031001004

World Medical Association. (2000). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA: The Journal of the American Medical Association, 284*(23), 3043-3045.

Ximénez, C., Maydeu-Olivares, A., Shi, D., & Revuelta, J. (2022). Assessing cutoff values of SEM fit indices: Advantages of the unbiased SRMR index and its cutoff criterion based on communality. *Structural Equation Modeling: A Multidisciplinary Journal, 29*(3), 368-380. https://doi.org/10.1080/10705511.2021.1992596

Yoder, N. (2014). *Self-assessing social and emotional instruction and competencies: A tool for teachers.* Center on Great Teachers and Leaders at American Institutes for Research.

Appendix 1 - Items of the SECTRS-14 (Grazzani, Martinsone, Simoes, Cavioni & Pepe, 2023).

The statements below describe your thoughts, feelings, and actions in the classroom and in situations at your school. For each item, please indicate the extent to which you agree or disagree with the statement provided. There are no right or wrong answers, so please be as honest as possible.						
	Strongly Disagree = 1	Disagree = 2	Somewhat Disagree= 3	Somewhat agree = 4	Agree = 5	Strongly Disagree = 6
1) I have a close relationship with my students.						
2) I nearly always stay calm when a student upsets me.						
3) My students' safety is an important factor in the decisions I make.						
4) Staff members seek my advice when resolving a problem.						
5) I am aware of how all of my students are feeling.						
6) I remain calm when addressing student misbehavior.						
7) I problem-solve with students when there is a problem or argument.						
8) I am good at understanding how my students' feel.						
9) Staff members at my school respect me.						
10) I create a sense of community in my classroom.						
11) I am able to manage my emotions and feelings in healthy ways.						
12) I consider my students' well-being when making decisions.						
13) I pay attention to the emotions of staff members at my school.						
14) I build positive relationships with my students' families.						