frontiers | Frontiers in Education

TYPE Brief Research Report
PUBLISHED 07 January 2025
DOI 10.3389/feduc.2024.1519327

OPEN ACCESS

EDITED BY
Athanasios Mogias,
Democritus University of Thrace, Greece

REVIEWED BY
Natanael Karjanto,
Sungkyunkwan University, Republic of Korea
Theodora Boubonari,
Democritus University of Thrace, Greece

*CORRESPONDENCE
Ivana Stepanović Ilić
✉ istepano@f.bg.ac.rs

RECEIVED 29 October 2024
ACCEPTED 23 December 2024
PUBLISHED 07 January 2025

CITATION
Stepanović Ilić I, Peixoto F, Dabić Boričić M, Videnović M, Castro Silva J, Ferreira N and Monteiro V (2025) Construction of a scale measuring primary teachers' self-efficacy to teach STEAM: the STEAM-TSES.
Front. Educ. 9:1519327.
doi: 10.3389/feduc.2024.1519327

Construction of a scale measuring primary teachers' self-efficacy to teach STEAM: the STEAM-TSES

Ivana Stepanović Ilić[1]*, Francisco Peixoto[2], Milana Dabić Boričić[3], Marina Videnović[4], Jose Castro Silva[2], Nadia Ferreira[2] and Vera Monteiro[2]

[1]Department of Psychology, Faculty of Philosophy, University of Belgrade, Belgrade, Serbia, [2]ISPA – University Institute, Centre for Research in Education (CIE-ISPA), Lisbon, Portugal, [3]Faculty of Teacher Education, University of Belgrade, Belgrade, Serbia, [4]Institute of Psychology, Faculty of Philosophy, University of Belgrade, Belgrade, Serbia

The STEAM-TSES was developed to measure teachers' self-efficacy (TSE) in teaching STEAM disciplines due to the lack of instruments measuring specific TSE dimensions, especially in STEAM education. This paper presents the validation results based on a sample of 805 Portuguese and Serbian primary school teachers. The scale comprises 36 items measuring TSE related to mathematics, biology, chemistry, science, and arts, the ability to motivate students, and confidence in integrating STEAM disciplines. The measurement model tested through Exploratory Structural Equation Modeling and Confirmatory Factor Analysis showed an acceptable fit to the data, supporting the existence of nine factors: five related to TSE in teaching specific subjects, three focused on motivating students in mathematics, science, and arts, and one related to integrated teaching. The analyses show good psychometric characteristics of the STEAM-TSES, and the correlations with the TSE scale indicate the validity of distinguishing constructs associated with STEAM education.

KEYWORDS

teachers' self-efficacy, STEAM, primary school, scale development, scale validation

1 Introduction

The concept of TSE is rooted in Rotter's considerations of locus of control and Bandura's social cognitive theory (Perera et al., 2019; Skaalvik and Skaalvik, 2007). Bandura (1986, 1997) triadic reciprocity model emphasizes the interplay of behavior, cognitive, emotional, and environmental factors, viewing humans as self-aware agents capable of self-regulation. Self-efficacy, central to this theory, reflects beliefs in one's ability to accomplish tasks, shaping perceptions of opportunities, challenges, effort, and persistence (Fernandez et al., 2016; Skaalvik and Skaalvik, 2010). TSE specifically refers to teachers' beliefs in their capacity to plan, manage, and implement activities necessary to achieve educational goals (Skaalvik and Skaalvik, 2010). Research shows that higher TSE positively impacts teachers and students (Lazarides and Warner, 2020; Perera et al., 2019; Zee and Koomen, 2016). Tschannen-Moran and Woolfolk Hoy (2001) found that higher TSE is associated with more effective planning, innovative teaching, and coping with challenges, leading to greater job satisfaction. Recent reviews further highlight TSE's significant influence on teacher well-being and engagement, which constructively facilitates students' achievements and self-efficacy (Fernandez et al., 2016; Klassen and Tze, 2014; Zee and Koomen, 2016).

Scholars (Lazarides and Warner, 2020; Perera et al., 2019) emphasize the contextual nature of self-efficacy in Bandura (1997) theory, emphasizing the various domains where TSE can manifest and be measured. The most influential model of general TSE is three-dimensional, encompassing teachers' perceptions of their instructional strategies, influence on student engagement, and classroom management abilities (Tschannen-Moran and Woolfolk Hoy, 2001). However, Klassen et al. (2011) warn that measures of TSE should meet the "right balance between specificity and generality" and that teacher's SE across different subject domains is often overlooked.

In our opinion such domains are related to the Science, Technology, Engineering, Arts and Math (STEAM) education which has garnered significant attention in recent decades, demonstrating improvements in students' critical thinking, creativity, collaboration, and problem-solving skills (Perignat and Katz-Buonincontro, 2019). These skills are essential for individuals in a knowledge-driven world and preparing for future work tasks (OECD, 2024; Oliveri et al., 2017). However, successful implementation of STEAM requires skilled teachers, especially in early schooling, given students' ongoing development in experimentation and reflection capacities (Bassachs et al., 2020; Silva-Hormazábal and Alsina, 2023).

In STEAM education studies, TSE is often considered a dependent variable within the pre-test/post-test design when evaluating teacher development programs aimed at implementing STEAM concepts (Romero-Ariza et al., 2021; Jiang et al., 2024). For that purpose, some researchers have employed general TSE scales focused on the previously mentioned three domains, which are not specifically aligned with STEAM (e.g., Boice et al., 2021). Alternatively, *ad hoc* relatively short scales have been constructed, including other aspects associated with teaching STEAM, such as beliefs or perceptions about it and enjoyment in applying the STEAM approach (DeJarnette, 2018; Jamil et al., 2018). In some cases, scales focused on pedagogical practices related to STEAM, such as inquiry-based learning, have been used (e.g., Romero-Ariza et al., 2021). The scale developed by the Friday Institute for Educational Innovation (2012) has frequently been adapted and used to investigate teachers' perceptions of STEM or to evaluate professional development programs (e.g., Çiftçi et al., 2020; Estévez-Mauriz and Baelo, 2021). That scale assesses teachers' learning and leadership attitudes; TSE and teaching outcomes expectancy related to mathematics; STEM career awareness; and teachers' perception of students' behavior related to STEM education. Researchers (e.g., Jamil et al., 2018) suggest that general TSE scales may miss essential confidence factors specific to subjects, particularly in the integration of arts and sciences crucial to STEAM. They argue that teachers need specialized scales that capture nuances in motivation and confidence across disciplines, as each component of STEAM demands unique pedagogical approaches.

Perignat and Katz-Buonincontro (2019) highlight that teachers need robust self-efficacy to effectively foster students' interdisciplinary skills, underscoring the importance of developing more specialized TSE scales. Moreover, research indicates that educators across various countries and educational systems (e.g., Çiftçi et al., 2020) encounter differing levels of STEAM training. This diversity suggests that self-efficacy scales should be flexible and responsive to distinct educational contexts and the unique challenges teachers encounter.

Considering the aspects outlined above, our objective was to develop an instrument specifically designed to measure TSE within the context of STEAM education and its associated disciplines. While some studies have explored TSE in STEM (e.g., Yang et al., 2023), few have addressed STEAM, particularly at the primary school level (e.g., Jamil et al., 2018). Given the unique challenges of teaching STEAM and the lack of instruments, measuring TSE in this area among primary educators is critical. Thus, research aims of this study are to present the development of a scale assessing teachers' self-efficacy (TSE) in teaching STEAM disciplines and to validate its psychometric characteristics, on a sample of primary school educators from Serbia and Portugal. These countries were selected as examples of different education policies related to the STEAM approach. STEAM education is essential for equipping students with transversal skills across various fields, preparing them to address the complex and unpredictable challenges of the future (European Schoolnet, 2018). In Portugal, in line with this orientation of EU educational policy, curriculum guidelines emphasize an integrated, multi- and interdisciplinary approach. In contrast, Serbia does not explicitly recognize the STEM/STEAM methodology. While similar initiatives exist, they are underdeveloped and not systematically implemented. Although the law and education strategy stress the importance of cross-curricular competencies in primary education (Law on the Foundations of the Education System, 2018), the curriculum does not mandate interdisciplinary classes.

2 Method

2.1 Participants and procedures

Participants were 805 primary school teachers (75.5% from Serbia: 90% female, and 24.5% from Portugal: 93% female) instructing 7–10 years old students, who responded an anonymous online survey voluntarily. Ethical approval was obtained from both university institutions. Mean age (Serbia: *M* - 47.5, *SD* - 8.7; Portugal: *M* - 48.6, *SD* - 7.9) and teaching experience (Serbia: *M* - 21, *SD* - 10.2; Portugal: *M* – 23.8, *SD* – 10.2) were similar in both countries. Portuguese teachers had more training related to STEAM (36.7% of the national sample size) or integrated approaches to teaching (69.9%) than Serbian teachers (STEAM - 10%, integrated teaching - 24.3%). The majority of teachers have a bachelor's or master's educational level (Serbia: 29.7% bachelor, 59.3% master; Portugal: 82.7% bachelor, 14.7% master).

2.2 Development of a scale measuring TSE in STEAM context

The scale was designed based on the widely used Science Teaching Efficacy Belief Scale (Enochs and Riggs, 1990) and the scale assessing TSE in mathematics developed by Friday Institute for Educational Innovation (2012). The objective was to gauge teachers' confidence in teaching school subjects related to STEAM disciplines (mathematics, arts, and science domains), as well as their ability to involve students' learning in these areas and integrate STEAM disciplines effectively.

The scale was structured into three items' groups: the first assesses TSE in specific STEAM subjects, the second focuses on motivating students in STEAM disciplines, and the third evaluates confidence in integrating them into teaching. This led to nine subscales: (1) TSE in teaching mathematics; (2) TSE in teaching biology; (3) TSE in teaching chemistry; (4) TSE in teaching physics; (5) TSE in teaching arts; (6) TSE in motivating students to learn mathematics; (7) TSE in motivating students in science; (8) TSE in motivating students in arts; and (9) TSE in integrating STEAM disciplines.

Next, the items were translated into Serbian and Portuguese by bilingual experts familiar with both educational and psychological terminology, followed by back-translation to check on the translation's fidelity to the original meaning. Any differences between the original and back-translated items were carefully reviewed and minimal discrepancies were found. Local teachers reviewed the translated items to confirm clarity and relevance within each cultural context. Three teachers from each country provided feedback into how items might be interpreted in real classroom settings, and their suggestions were used to refine the items' accuracy.

2.3 Measures

2.3.1 STEAM-teacher self-efficacy scale (STEAM-TSES)

The STEAM-TSE scale comprised 36 items aggregated in nine dimensions. Five dimensions with 20 items measure TSE within a specific subject domain related to STEAM: mathematics (e.g., *I am confident that I can explore mathematics with students in a way that they understand the meaning behind mathematics notation and facts*), biology (e.g., *I am confident that I can teach biology concepts effectively*), chemistry (e.g., *I am confident that I can answer students' questions about chemistry*), physics (e.g., *I am confident that I can run physics experiments*) and arts (e.g., *I am confident that I can teach arts/artistic expression effectively*). Three dimensions with 12 items (four per one dimension) assess TSE in motivating students to learn three STEAM disciplines for these topics: mathematics (e.g., *I am confident that I'm able to increase students' interest in mathematics*), science (e.g., *I am confident that I can motivate students who show low interest in science*) and arts (e.g., *I am confident that I can help students to value learning about the arts/artistic expression concepts*). The last dimensions with four items measure TSE in integrating STEAM disciplines (e.g., *I feel confident in combining different disciplines to teach about particular phenomena,* e.g., *when explaining the cycle of water in nature to combine concepts as evaporating, living beings, condensation, temperature, etc.*). The 7-point Likert scale was administered, offering two options: teachers indicated their confidence about a particular item (1 – Not confident at all | 7 - Absolutely confident) or agreement with an item (1 – Completely Disagree | 7 – Completely Agree).

2.3.2 Global teachers' self-efficacy scale

To assess global TSE, we used a scale adapted by Peixoto et al. (2018), based on Morgan's (2011) general TSE scale. This scale comprises nine items tapping teaching strategies (e.g., *I am confident that I can teach all the subjects on the school curriculum effectively*) and classroom management (e.g., *I am confident that I can manage inappropriate behavior in the class*). As for the STEAM-TSES, teachers answered on a 7-point Likert scale.

2.4 Data analysis

The psychometric properties of the STEAM-TSES were analyzed by assessing its structure, reliability, discriminant, and convergent validity. To examine the scale's structure, we conducted Exploratory Structural Equation Modeling (ESEM) and Confirmatory Factor Analyses (CFA) in MPlus (v. 8.0, Muthén and Muthén, 1998–2016) using the Maximum Likelihood Robust (MLR) estimator. These are standard statistical techniques recommended for identifying latent factors behind observable variables measuring phenomena such as beliefs (Schreiber et al., 2006). The sample was randomly split in two, with ESEM applied to one half and CFA to the other. Model fit was evaluated using the following indices: Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). CFI and TLI values above 0.90 indicate a good fit, and values above 0.95 are considered very good. The RMSEA and SRMR, values below 0.08 suggest a good fit, along with a narrow confidence interval (Hu and Bentler, 1999; Schumacker and Lomax, 2004).

Reliability was assessed using Cronbach's alpha, McDonald's omega, and composite reliability. Convergent and discriminant validity were tested following Cheung and Chang's (2017) and Fornell and Larcker's (1981) guidelines. Cross-country measurement invariance was tested using the differences in CFI and RMSEA due to the sensitivity of the chi-square test to sample size. Differences in CFI less than 0.01 and in RMSEA less than 0.015 were considered indicative of measurement invariance (Chen, 2007).

3 Results

The Exploratory Structural Equation Modeling (ESEM) conducted on the 36 items of the scale shows an acceptable fit to the data, χ^2 (342) = 778.13, $p < 0.001$, CFI = 0.955, TLI = 0.918, RMSEA = 0.056, [0.051, 0.061], SRMR = 0.013. The results reveal that the majority of items are grouped around the same factor. However, item 35, which was expected to load on the dimension related to TSE in motivating students for mathematics, showed a higher factor loading in the factor associated with items related to mathematics teaching. Additionally, items from the Arts Teaching dimension loaded higher on a combined factor that included both Arts Teaching and Arts Motivation items, rather than on the factor specifically associated with Arts Teaching. All other items had factor loadings higher than 0.45 on a single factor, and all items grouped congruently with those of the same content (see Supplementary material). Following this, we carried out a Confirmatory Factor Analysis (CFA) on the other half of the sample, testing two models: one with nine factors corresponding to the originally conceptualized dimensions, and another with eight factors, merging the Arts Teaching and Arts Motivation items into a single factor based on the ESEM results. The results showed that the 9-factor solution better fit the data than the 8-factor solution (Table 1). Considering the whole sample, the 9-factor solution fits adequately the data (Table 1).

Due to the high correlations between Math Teaching and Math Motivation, Arts Teaching and Arts Motivation, and Chemistry Teaching and Physics Teaching (Table 2), we tested different models merging these dimensions, including all participants in the study.

Firstly, we tested a model combining Arts Teaching with Arts Motivation (Table 1, 8-factor model A), followed by the second model combining Math Teaching with Math Motivation, the third combining Chemistry Teaching with Physics Teaching, and finally, the model integrating all these combinations (a 6-factor model, aggregating Arts

TABLE 1 Goodness-of-fit indices for the tested models.

	c^2	df	*p*	CFI	TLI	RMSEA	SRMR	AIC	BIC
9-factor model (half sample)	1324.96	558	<0.001	0.918	0.908	0.058, [0.054, 0.063]	0.045	30554.63	31130.13
8-factor model A (half sample)	1481.91	566	<0.001	0.902	0.891	0.063, [0.060, 0.067]	0.047	30763.7	31307.22
9-factor model (all sample)	1599.73	558	<0.001	0.939	0.931	0.048, [0.045, 0.051]	0.036	59766.75	60442.24
8-factor model A (all sample)	1912.87	566	<0.001	0.921	0.912	0.054, [0.052, 0.057]	0.04	60246.29	60884.25
8-factor model B (all sample)	1797.75	566	<0.001	0.928	0.92	0.052, [0.049, 0.055]	0.037	60071.54	60709.49
8-factor model C (all sample)	2237.86	566	<0.001	0.902	0.891	0.061, [0.058, 0.063]	0.035	60741.83	61379.78
6-factor model (all sample)	2727.4	579	<0.001	0.874	0.863	0.068, [0.065, 0.063]	0.04	61524.96	62101.94

The 8-factor model A combines the arts teaching and arts motivation dimensions into a single factor. The 8-factor model B combines the math teaching and math motivation items, and the 8-factor model C combines the physics and chemistry teaching items. The 6-factor model includes all these aggregations from each 8-factor model in a single model.

TABLE 2 Correlation matrix between the dimensions of the STEAM-TSES, correlation with TSE and AVE for each dimension.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	TSE	AVE
F1		0.46	0.46	**0.85**	0.54	0.41	0.43	0.60	0.57	0.63	0.77
F2	0.21		**0.89**	0.46	0.56	0.48	0.45	0.56	0.55	0.44	0.82
F3	0.22	**0.80**		0.49	0.52	0.41	0.40	0.59	0.55	0.44	0.79
F4	0.73	0.21	0.24		0.55	0.40	0.43	0.66	0.55	0.64	0.73
F5	0.29	0.31	0.27	0.31		0.65	0.67	**0.72**	0.59	0.41	0.78
F6	0.16	0.23	0.16	0.16	0.43		**0.82**	0.62	0.53	0.32	0.81
F7	0.18	0.20	0.16	0.18	0.45	0.68		0.66	0.54	0.33	0.79
F8	0.35	0.32	0.35	0.43	0.52	0.39	0.43		**0.71**	0.48	0.77
F9	0.32	0.30	0.30	0.30	0.34	0.28	0.30	0.51		0.43	0.75

Values above the diagonal are correlations among constructs, and values below the diagonal are squared correlations. All correlation values are statistically significant at $p < 0.001$. In bold: correlations higher than 0.70. F1– math teaching, F2 – arts teaching, F3 – arts motivation, F4 – math motivation, F5 – biology teaching, F6 – chemistry teaching, F7 – physics teaching, F8 – science motivation, F9 – integrated teaching. TSE – teacher self-efficacy (Global), AVE – average variance extracted.

Teaching with Arts Motivation, Math Teaching Math Motivation, and Chemistry Teaching with Physics Teaching). The results (Table 1) indicate that none of these models improved the goodness-of-fit. The AIC and BIC values of the 9-factor model were the lowest compared to the 8-factor and 6-factor models and the other fit indices were also noticeably better (see Table 1).

All dimensions showed very good reliability, presenting scores above 0.90 in all dimensions for the different indicators of reliability used (Table 3). Convergent validity for each dimension was assessed considering that a measure has good convergent validity when the average variance extracted (AVE) is higher than 0.50 (Fornell and Larcker, 1981), the factor loadings are not significantly lower than 0.50 and the composite reliability (CR) is higher than 0.70 (Cheung and Chang, 2017). Analyses of the AVE (Table 2), the factor loadings, and the CR of the scale dimensions (Table 3) showed good convergent validity for all dimensions.

To test discriminant validity, Fornell and Larcker (1981) proposed comparing the AVE with the squared correlation between constructs. If the AVE for both factors is higher than the squared correlation between them, discriminant validity is established. Cheung and Chang (2017) also suggested that correlations between constructs should not exceed 0.70. Following Fornell and Larcker's (1981) guidelines, all dimensions, except for the pair Arts Teaching and Arts Motivation, demonstrate good discriminant validity as the AVE is higher than the squared correlation between the factors. However, applying Cheung and Chang's (2017) criterion, five pairs of factors present discriminant validity issues. These pairs are Math Teaching/Math Motivation, Arts Teaching/Arts Motivation, Chemistry Teaching/Physics Teaching, Biology Teaching/Science Motivation, and Science Motivation/Integrated Teaching, as the correlations between the factors within each pair exceed 0.70 (Table 2).

The correlations between the STEAM-TSES dimensions and the global TSE (Table 2) showed moderate and positive relationships. The highest correlations are with the Math Teaching and Math Motivation dimensions, while the weakest are with Physics and Chemistry Teaching.

To test measurement invariance, we sequentially tested models of configural, metric and scalar invariance. The results (Table 4) indicate that we have scalar measurement invariance, implying that we can compare the means obtained with this scale across the two countries.

TABLE 3 Descriptive statistics, reliability, and lowest factor loading for each dimension of the STEAM-TSES.

	M	SD	$M_{ri(t-i)}$[a]	CR	Cronbach's α	McDonald's ω	Lowest factor loading
F1. Math teaching	5.95	0.85	0.77	0.98	0.90	0.90	0.79
F2. Arts teaching	5.92	0.78	0.84	0.97	0.93	0.93	0.86
F3. Arts motivation	6.02	0.82	0.90	0.96	0.96	0.96	0.92
F4. Math motivation	5.62	1.00	0.80	0.92	0.91	0.91	0.83
F5. Biology teaching	5.90	0.84	0.84	0.93	0.93	0.93	0.84
F6. Chemistry teaching	5.79	1.05	0.86	0.97	0.94	0.94	0.80
F7. Physics teaching	4.83	1.51	0.86	0.97	0.94	0.94	0.82
F8. Science motivation	5.00	1.40	0.83	0.93	0.93	0.93	0.86
F9. Integrated teaching	5.61	1.01	0.82	0.92	0.92	0.92	0.83

CR, composite reliability. [a]M of part-whole corrected item-total correlations.

4 Discussion

Previous research has accentuated the importance of measuring TSE, indicating its positive impact on teachers and their pedagogical strategies (Perera et al., 2019; Tschannen-Moran and Woolfolk Hoy, 2001; Zee and Koomen, 2016), as well as on students' outcomes (Fernandez et al., 2016; Klassen and Tze, 2014; Zee and Koomen, 2016). However, there is a lack of instruments measuring TSE in the context of the increasingly prevalent STEAM education, particularly at the primary school level. Thus, this study aims to develop a STEAM self-efficacy scale for primary school teachers: the STEAM-TSES.

The STEAM approach requires integrating different subject areas, making it essential to find an optimal level of specificity for TSE measures. TSE may vary across domains (Klassen et al., 2011; Perera et al., 2019), especially considering differences in pre-service training and school curricula. Teaching science demands theoretical and methodological knowledge from various domains (chemistry, physics, biology), and TSE can differ in these areas, particularly for primary school pupils, whose understanding of scientific concepts is limited by cognitive developmental characteristics (e.g., Bassachs et al., 2020). Thus, our aim was to assess TSE not only in mathematics but also in science and the arts, and to create a subscale dedicated to the integration of all domains.

The analyses conducted suggest that the scale exhibits very satisfactory psychometric properties. It comprises 36 items aggregated into 9 distinct factors, despite some discriminant validity issues. As previously noted, the dimensions of Math Teaching/Math Motivation, Arts Teaching/Arts Motivation and Chemistry Teaching/Physics Teaching showed very high correlations. Nevertheless, we tested three alternative eight-factor models and a six-factor model, all showing a worse fit than the nine-factor solution. Consequently, to avoid multicollinearity problems, we recommend caution when using Math Teaching/Math Motivation, Arts Teaching/Arts Motivation and Chemistry Teaching/Physics Teaching dimensions simultaneously in structural equation modeling or other regression-based analyses. Our results justify the measurement of TSE across different science domains. They indicate that even though TSE is similar across nine dimensions, these dimensions are not identical. This instrument may be particularly useful considering the differences in educational systems and teachers' previous experience, which can affect TSE in various subjects.

The need for a specialized STEAM-TSES is supported by moderate and positive correlations between the nine STEAM-TSES dimensions and the global measure of TSE. The highest correlations are found among closely related dimensions, such as Math teaching and Math motivation, Arts teaching and Arts motivation, and Physics and Chemistry teaching. These correlations are expected, as they reflect specific aspects of TSE that collectively contribute to a broader sense of competence, often referred to as global TSE. Furthermore, the moderate to strong correlations between the STEAM-TSES dimensions and the measure of global TSE reinforce the idea that these dimensions play a key role in shaping an overall sense of teaching competence. However, the AVE greater than 0.5 indicates that while the STEAM-TSES dimensions and the global TSE measure are related, they are not identical constructs. As expected, the highest correlations were found between the global TSE and dimensions related to mathematics teaching. This finding further suggests that the global TSE does not sufficiently capture other aspects of STEAM teaching practices that are more distant from mathematics, such as the arts, science disciplines, and the integration of different subjects.

The STEAM-TSES demonstrated very good reliability and convergent validity. High-reliability scores suggest the potential for reducing the number of items per dimension without compromising reliability. The results of the measurement invariance analyses revealed the scale's utility across both countries. Moreover, with scalar invariance, the means obtained in each country can be reliably compared.

We believe the STEAM-TSES scale is relevant for both countries, though its cultural significance varies due to the differing levels of integration and support for STEAM education in each. In Portugal, the scale could be used as a measuring tool of TSE in teaching and motivating students within STEAM disciplines, as well as their confidence in applying the integrated, interdisciplinary approach central to the curriculum. In Serbia, it offers valuable insights into TSE in STEAM, particularly in adapting traditional teaching methods to incorporate interdisciplinary elements, and highlights areas where the educational system needs to strengthen STEAM implementation.

Given the differing European regions and similar yet distinct curricula of these two countries, the STEAM-TSE scale may have wider applicability across different countries. As previously mentioned, various educational contexts can benefit from using this scale, but the interpretation of results should be sensitive to the varying degrees of policy support for STEAM in particular country.

Initially aimed at measuring TSE in early primary education, we believe it could also serve teachers instructing older grades. General item formulation supports its potential use at middle and

TABLE 4 Test of measurement invariance.

	c^2	*df*	*p*	CFI	TLI	RMSEA	D CFI	D RMSEA
Configural	2557.4	1,116	<0.001	0.927	0.918	0.057, [0.054,0.060]	-	-
Metric	2566.5	1,143	<0.001	0.928	0.921	0.056, [0.053,0.058]	−0.001	0.001
Scalar	2644.9	1,170	<0.001	0.926	0.92	0.056, [0.053,0.059]	0.002	0

secondary levels. Nonetheless, these assumptions necessitate further research, involving administering the scale in various countries and at different educational levels.

STEAM-TSES can be a useful tool for teacher training programs, curriculum designers, and policymakers. The STEAM-TSES allows teacher training programs to assess teachers' confidence in specific STEAM disciplines. Data gathered using the scale can guide trainers in providing individualized support to pre-service and experienced teachers. Curriculum designers can use insights from the scale to create STEAM resources that align with teachers' existing strengths while supporting teacher perceived low self-efficacy. Additionally, school administrators can use the STEAM-TSES scale to track the effectiveness of initiatives aimed at improving teacher efficacy in STEAM disciplines.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ISI: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing. FP: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MD: Conceptualization, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing. MV: Conceptualization, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing. JC: Conceptualization, Investigation, Methodology, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing. NF: Conceptualization, Investigation, Project administration, Writing – original draft, Writing – review & editing. VM: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Portuguese Science and Technology Foundation (FCT) under grants [UIDP/04853/2020; doi:10.54499/UIDB/04853/2020] and [UIDB/04853/2020; doi:10.54499/UIDP/04853/2020] awarded to CIE-ISPA. Ministry of Science, Technological Development, and Innovation of the Republic of Serbia as part of the funding for scientific research work at the University of Belgrade - Faculty of Philosophy (contract number 451–03-66/2024–03/200163).

Conflict of interest

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

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1519327/full#supplementary-material

References

Bandura, A. (1986). Social foundations of thought and action. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). Self-efficacy. New York: Freeman.

Bassachs, M., Cañabate, D., Nogué, L., Serra, T., Bubnys, R., and Colomer, J. (2020). Fostering critical reflection in primary education through STEAM approaches. *Educ. Sci.* 10:384. doi: 10.3390/educsci10120384

Boice, K. L., Jackson, J. R., Alemdar, M., Rao, A. E., Grossman, S., and Usselman, M. (2021). Supporting teachers on their STEAM journey: a collaborative STEAM teacher training program. *Educ. Sci.* 11:105. doi: 10.3390/educsci11030105

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14, 464–504. doi: 10.1080/10705510701301834

Cheung, G. W., and Chang, W. (2017). Current approaches for assessing convergent and discriminant validity with SEM: issues and solutions. *Acad. Manag. Proceed.* 2017:12706. doi: 10.5465/AMBPP.2017.12706abstract

Çiftçi, A., Topçu, M. S., and Foulk, J. A. (2020). Pre-service early childhood teachers' views on STEM education and their STEM teaching practices. *Res. Sci. Technol. Educ.* 40, 207–233. doi: 10.1080/02635143.2020.1784125

DeJarnette, N. K. (2018). Implementing STEAM in the early childhood classroom. *Eur. J. STEM Educ.* 3:18. doi: 10.20897/ejsteme/3878

Enochs, L. G., and Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: a preservice elementary scale. *Sch. Sci. Math.* 90, 694–706. doi: 10.1111/j.1949-8594.1990.tb12048.x

Estévez-Mauriz, L., and Baelo, R. (2021). How to evaluate the STEM curriculum in Spain? *Mathematics* 9:236. doi: 10.3390/math9030236

European Schoolnet (2018). Science, technology, engineering and mathematics education policies in Europe. Scientix observatory report. Brussels: European Schoolnet. Available at: https://www.scientix.eu/documents/10137/782005/Scientix_Texas-Instruments_STEM-policies-October-2018.pdf/d56db8e4-cef1-4480-a420-1107bae513d5 (Accessed October 15, 2024).

Fernandez, A. P. O., Ramos, M. F. H., Silva, S. S. C., Nina, K. C. F., and Pontes, F. A. R. (2016). Overview of research on teacher self-efficacy in social cognitive perspective. *Anales de Psicología* 32, 793–802. doi: 10.6018/analesps.32.3.220171

Fornell, C., and Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *J. Mark. Res.* 18, 382–388. doi: 10.1177/002224378101800313

Friday Institute for Educational Innovation (2012). Teacher efficacy and attitudes toward STEM survey: Elementary teachers. Raleigh: Friday Institute for Educational Innovation.

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling: A Multidiscip. Res.* 6, 1–55. doi: 10.1080/10705519909540118

Jamil, F. M., Linder, S. M., and Stegelin, D. A. (2018). Early childhood teacher beliefs about STEAM education after a professional development conference. *Early Childhood Educ. J.* 46, 409–417. doi: 10.1007/s10643-017-0875-5

Jiang, H., Chugh, R., Zhai, X., Wang, K., and Wang, X. (2024). Longitudinal analysis of teacher self-efficacy evolution during a STEAM professional development program: a qualitative case study. *Human. Soci. Sci. Commun.* 11, 1–16. doi: 10.1057/s41599-024-03655-5

Klassen, R. M., and Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: a meta-analysis. *Educ. Res. Rev.* 12, 59–76. doi: 10.1016/j.edurev.2014.06.001

Klassen, R. M., Tze, V. M., Betts, S. M., and Gordon, K. A. (2011). Teacher efficacy research 1998–2009: signs of progress or unfulfilled promise? *Educ. Psychol. Rev.* 23, 21–43. doi: 10.1007/s10648-010-9141-8

Law on the Foundations of the Education System. (2018). The Official Gazette of the Republic of Serbia, No. 88/2017, 27/2018. Available at: https://pravno-informacioni-sistem.rs/eli/rep/sgrs/skupstina/zakon/2017/88/1/reg (Accessed October 15, 2024).

Lazarides, R., and Warner, L. M. (2020). "Teacher self-efficacy" in Oxford research encyclopedia, Education. Ed. G. W. Noblit (Oxford: Oxford University Press), 1–22.

Morgan, M. (2011). Resilience and recurring adverse events: testing an assets-based model of beginning teachers' experiences. *Ir. J. Psychol.* 32, 92–104. doi: 10.1080/03033910.2011.613189

Muthén, L. K., and Muthén, B. O. (1998–2016). Mplus user's guide. *7th* Edn. Los Angeles: Muthén & Muthén.

OECD (2024). OECD contributions to the 2030 agenda and beyond: Shaping a sustainable future for all. Paris: OECD Publishing.

Oliveri, M. E., Lawless, R., and Molloy, H. (2017). A literature review on collaborative problem solving for college and workforce readiness. *ETS Res. Rep. Ser.* 2017, 1–27. doi: 10.1002/ets2.12133

Peixoto, F., Wosnitza, M., Pipa, J., Morgan, M., and Cefai, C. (2018). "A multidimensional view on pre-service teacher resilience in Germany, Ireland, Malta and Portugal" in Resilience in education: Concepts, contexts and connections. eds. A. Wosnitza, F. Peixoto, S. Beltman and C. F. Mansfield (Cham: Springer), 73–89.

Perera, H. N., Calkins, C., and Part, R. (2019). Teacher self-efficacy profiles: determinants, outcomes, and generalizability across teaching level. *Contemp. Educ. Psychol.* 58, 186–203. doi: 10.1016/j.cedpsych.2019.02.006

Perignat, E., and Katz-Buonincontro, J. (2019). STEAM in practice and research: an integrative literature review. *Think. Skills Creat.* 31, 31–43. doi: 10.1016/j.tsc.2018.10.002

Romero-Ariza, M., Quesada, A., Abril, A. M., and Cobo, C. (2021). Changing teachers' self-efficacy, beliefs and practices through STEAM teacher professional development. *J. Study Educ. Dev.* 44, 942–969. doi: 10.1080/02103702.2021.1926164

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *J. Educ. Res.* 99, 323–338. doi: 10.3200/JOER.99.6.323-338

Schumacker, R. E., and Lomax, R. G. (2004). A beginner's guide to structural equation modeling. *2nd* Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

Silva-Hormazábal, M., and Alsina, Á. (2023). Exploring the impact of integrated STEAM education in early childhood and primary education teachers. *Educ. Sci.* 13:842. doi: 10.3390/educsci13080842

Skaalvik, E. M., and Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *J. Educ. Psychol.* 99, 611–625. doi: 10.1037/0022-0663.99.3.611

Skaalvik, E. M., and Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: a study of relations. *Teach. Teach. Educ.* 26, 1059–1069. doi: 10.1016/j.tate.2009.11.001

Tschannen-Moran, M., and Woolfolk Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teach. Teach. Educ.* 17, 783–805. doi: 10.1016/S0742-051X(01)00036-1

Yang, W., Wu, R., and Li, J. (2023). Development and validation of the STEM teaching self-efficacy scale (STSS) for early childhood teachers. *Curr. Psychol.* 42:7275. doi: 10.1007/s12144-021-02074-y

Zee, M., and Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: a synthesis of 40 years of research. *Rev. Educ. Res.* 86, 981–1015. doi: 10.3102/0034654315626801