MODESTUM
EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(6), em2647
ISSN:1305-8223 (online)
OPEN ACCESS
Research Paper
https://doi.org/10.29333/ejmste/16507

Self-efficacy in computational thinking: Preliminary analysis of pre-service teachers' perception through a Portuguese tool

Rita Neves Rodrigues [1,2,3*], Cecília Costa [2,3], Maryam Abbasi [1,4], Fernando Martins [1,5,6,7], Sónia Brito-Costa [1,5*]

[1] Polytechnic University of Coimbra, Coimbra, PORTUGAL
[2] University of Trás-os-Montes and Alto Douro, Vila Real, PORTUGAL
[3] Research Center for Didactics and Technology in Trainer Training, University of Aveiro, Aveiro, PORTUGAL
[4] Research Center for Natural Resources Environment and Society, Polytechnic University of Coimbra, Coimbra, PORTUGAL
[5] inED–Center for Research and Innovation in Education, Polytechnic University of Coimbra, Coimbra, PORTUGAL
[6] Instituto de Telecomunicações, Delegação da Covilhã, Covilhã, PORTUGAL
[7] SPRINT–Sport Physical Activity and Health Research & Innovation Center, Polytechnic University of Coimbra, Coimbra, PORTUGAL

Received 11 December 2024 ▪ Accepted 01 May 2025

Abstract
Computational thinking is a fundamental problem-solving skill and has been integrated into the curricula of several countries, including Portugal. The successful implementation of this integration depends on adequately training pre-service teachers to develop the pedagogical knowledge necessary to incorporate this skill into their practice. Analyzing pre-service teachers' perception of self-efficacy in computational thinking is essential, as it influences the success of training programs. This study aimed to achieve three objectives: (1) translate and adapt the computational thinking self-efficacy scale into Portuguese (stage 1), (2) validate the scale (stage 2), and (3) compare pre-service teachers' perceptions of self-efficacy in computational thinking (stage 3). A sample of 43 participants was used in stage 1 and 382 participants in stage 2 and stage 3. The findings demonstrated that the scale had strong temporal stability, good item homogeneity, and alignment with its original structure. The Portuguese version exhibited adequate psychometric properties, making it a valid and reliable instrument for analyzing pre-service teachers' perceptions of self-efficacy in computational thinking. The absence of significant differences in self-efficacy perceptions in computational thinking between bachelor's and master's pre-service teachers suggests a shared underexposure to the concept, reflecting a broader lack of structured instruction to foster self-efficacy in this domain.

Keywords: initial teacher training, computational thinking, self-efficacy, questionnaire validation, self-efficacy scale

INTRODUCTION

Computational thinking has become increasingly prominent and relevant in educational research (Gao & Hew, 2022). Although Papert (1980) was a pioneer in introducing the subject to the scientific community, it was Wing (2006) who encouraged the integration of this mathematical ability into the school curriculum, thus boosting research into this subject (Grover & Pea, 2013). The literature identifies computational thinking as a skill encompassing essential problem-solving competencies (Ausiku & Matthee, 2021). In recent years, several studies have emerged in various countries (Bower et al., 2017; Esteve-Mon et al., 2019; Rich et al., 2020) on the advantages of promoting the development of the dimensions of computational thinking skills in students. Several studies on the advantages of developing computational thinking skills have been conducted in

✉ ritanevesrodrigues@hotmail.com **(*Correspondence)** ✉ mcosta@utad.pt ✉ maryam.abbasi@ipc.pt ✉ fmlmartins@esec.pt
✉ sonya.b.costa@gmail.com **(*Correspondence)**

Contribution to the literature

- This article presents the Portuguese translation and adaptation of a computational thinking self-efficacy scale, which allows us to analyze pre-service teachers' perceptions of their self-efficacy in computational thinking during initial teacher training.
- The article analyses and compares the perceptions of self-efficacy in computational thinking among pre-service teachers, exploring potential differences between those enrolled in bachelor's and master's degree programs.
- This article addresses a gap identified in the research on computational thinking in Portugal: the lack of training and preparation for teachers to effectively integrate this ability into teaching. It provides a starting point for the development of training programs in initial teacher education.

countries such as Thailand (Pewkam & Chamrat, 2022), Ireland (Butler & Leahy, 2021), New Zealand (Macann & Carvalho, 2021), and Norway (Kravik et al., 2022). In the 2022/2023 school year, Portugal incorporated computational thinking into the educational curriculum, starting from the first year of elementary school. This integration was made as part of the mathematics curriculum (Ministério da Educação, 2021). The curriculum emphasizes five key dimensions of computational thinking: abstraction, decomposition, pattern recognition, algorithms, and debugging (Ministério da Educação, 2021).

For the integration of computational thinking skills into teaching to be successful and reach all students, it is essential that teachers and pre-service teachers have training that allows them to develop the didactic knowledge necessary to integrate this mathematical skill into their future practices (Butler & Leahy, 2021; Sun et al., 2023). However, a lack of teacher training and capacity building in computational thinking remains a key obstacle to integrating this skill into education in Portugal (Graça & Colaço, 2024; Pinheiro et al., 2023; Ramos et al., 2022).

Although limited scientific evidence links self-efficacy to computational thinking ability, several authors argue that self-efficacy is one of the most important factors in achieving success in a training program (Kukul & Karatas, 2019; Şen, 2023). According to Bandura (1986), self-efficacy refers to an individual's belief in their ability to organize and execute actions to achieve specific goals. It plays a fundamental role in human behavior by influencing the ability to face adversity in order to achieve specific goals (Waddington, 2023). High levels of self-efficacy are just as important as the skills required for the training to be carried out, as it increases predisposition, motivation and commitment to the new task (Avcı & Deniz, 2022; Bandura, 1986). However, low levels of teacher self-efficacy negatively impact their ability to promote learning and effectively manage the classroom (Dong et al., 2024). Consequently, student motivation and success may be hindered, affecting their skills, cognitive development, and overall academic achievement (Durak et al., 2023).

According to Mason and Rich (2019), for a training program to effectively enhance self-efficacy perceptions, interventions must be continuous and long-term, incorporating both practical and reflective components throughout the training process. Similarly, Rodrigues et al. (2024) recommend that teacher training programs should include theoretical, practical, and reflective components, utilizing various data collection instruments to monitor not only the tasks performed by pre-service teachers but also their attitudes toward the development of computational thinking. In this context, in order to train pre-service teachers, it is important to analyze their perception of self-efficacy in computational thinking, as this can directly influence the successful implementation of training programs in this area (Şen, 2023). Similarly, analyzing pre-service teachers' perception of self-efficacy in computational thinking after the conclusion of a training program is essential to determine the success of that program (Avcı & Deniz, 2022). Therefore, analyzing pre-service teachers' perception of self-efficacy in computational thinking evaluates the effectiveness of the training program developed. It also ensures that pre-service teachers feel empowered and confident to effectively implement computational thinking skills in their classrooms (Bower & Falkner, 2015).

To evaluate implemented training, it is essential to use a valid and reliable instrument both before and after the intervention to analyze the evolution of pre-service teachers' perceptions of self-efficacy in computational thinking throughout the training (Avcı & Deniz, 2022; Román-González et al., 2017).

At the time this study was conducted, no valid instruments were identified for the Portuguese population to assess pre-service teachers' perceptions of self-efficacy in computational thinking. Therefore, the computational thinking self-efficacy scale developed by Kukul and Karatas (2019) was translated, adapted, and applied to pre-service teachers in initial teacher training programs in Portugal, including those pursuing bachelor's and master's degrees, with the aim of validating it in Portuguese.

A study conducted by Çakir et al. (2021) identified differences in pre-service teachers' perceptions of their competence in developing computational thinking skills between those in their first year of university and those in their fourth year. Similarly, a study conducted by Avcı and Deniz (2022) found differences in pre-service teachers' perceptions of self-efficacy in computational thinking based on their professional experiences, such as internships, and the levels of education they attended. These authors also state that, for pre-service teachers' self-efficacy levels to increase, the training they undergo must aim to enhance their understanding of computational thinking. Therefore, it is important to investigate whether the mere accumulation of professional experience by pre-service teachers is sufficient to promote an increase in their self-efficacy perception, or if, on the other hand, the implementation of specialized training programs is essential for the proper development of these competencies.

In Portugal, initial teacher training involves completing two degrees: a bachelor's degree and a master's degree. Enrolment in the second degree, the master's program, includes professional internships, providing pre-service teachers with more in-depth practical experience compared to those enrolled in the bachelor's program. Therefore, it is important to assess whether there are differences and specific needs between these two groups, which have distinct professional experiences.

Therefore, the following research questions emerge:

1. Is there a relationship between the dimensions of the self-efficacy scale and each level of education (undergraduate and graduate)?
2. Are there statistically significant differences in the perception of self-efficacy in computational thinking between pre-service teachers attending a bachelor's degree and those attending a master's degree?

The primary objective of this study is to address the lack of instruments in Portugal for assessing pre-service teachers' perceptions of self-efficacy in computational thinking. The existence of an instrument that allows for the analysis of pre-service teachers' perceptions of self-efficacy in computational thinking will help determine whether the training provided in initial teacher training is effective. This can be assessed by examining whether there are statistically significant differences between the first (bachelor's degree) and the final stage (master's degree) of the training. Additionally, the use of this instrument before and after a planned training program aimed at enhancing pre-service teachers' self-efficacy in computational thinking will enable the evaluation of the impact of this training. By bridging this gap, the study aims to support the integration of computational thinking skills into the Portuguese education system, particularly through the development of more targeted and effective training programs tailored to the specific needs of pre-service teachers.

In addition to contributing to the improvement of initial teacher training in Portugal, this study presents results that may be generalized to other educational contexts facing similar challenges in integrating computational thinking. With the increasing integration of computational thinking into curricula in various countries, understanding whether there are differences in self-efficacy perceptions among pre-service teachers at different levels of education, or based on their professional experiences, allows for reflection on the need to adapt training programs according to the characteristics of the target group.

To accomplish the defined objective, three stages were defined and will be presented in this paper. First, the computational thinking self-efficacy scale developed by Kukul and Karatas (2019) was translated and adapted to the Portuguese language and context. Second, the translated and adapted scale was validated. Finally, the study analyzed and compared the perceptions of self-efficacy in computational thinking between pre-service teachers pursuing a bachelor's degree and those enrolled in a master's degree.

MATERIALS AND METHODS

Study Design

This study aimed to translate, adapt, and validate the computational thinking self-efficacy scale for the Portuguese context and to analyze differences in self-efficacy perceptions among pre-service teachers. The design included three distinct stages:

(1) translation and adaptation stage,

(2) validation stage, and

(3) analysis and comparison stage.

The process of translation, adaptation, and validation of the scale was conducted in three stages. In stage 1, the initial translation and back-translation were performed by professional translators, followed by expert validation to ensure fidelity to the original instrument and cultural appropriateness. Temporal stability was determined using the intraclass correlation coefficient (ICC), and internal consistency (Cronbach's alpha) was evaluated for both the total scale and its dimensions to establish content validity. In stage 2, the internal consistency analysis (Cronbach's alpha) was repeated to assess preliminary construct validity. Then, the adequacy of the sample for factor analysis was verified using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity (BET). Then, an exploratory factor analysis was conducted using the principal component extraction method, with varimax rotation and Kaiser normalization. Determining the number of factors followed the Kaiser criterion, retaining only those with eigenvalues greater than 1.0, combined with the criterion that each factor should explain at least 5% of the total variance. The variance explained by the factors was analyzed, and the commonalities were examined, with items having factor loadings below 0.4 being excluded. The final retention of the factors was based on the theoretical interpretation of the extracted factors. Finally, in stage 3, the perceptions of groups at different levels of education (bachelor's and master's degrees) were compared using the independent samples t-test, calculating the effect size (Cohen's *d*), and the relationships between the dimensions of the scale were analyzed using Pearson's linear correlation.

Participants

Stage 1. Test-re-test sample

After translating and adapting the scale for the Portuguese language and context, a test-retest procedure was conducted to evaluate its reliability. This initial stage aimed to ensure the internal consistency and linguistic adequacy of the translated version before applying it to a larger sample for scale validation. This stage involved a sample of 43 participants from two universities in mainland Portugal, who completed the scale during the 2022/2023 academic year. The participants' average age was 21.67 years (standard deviation [SD] = 5.181), with ages ranging from 18 to 44 years, and all were female. The sample composition included 76.7% bachelor's degree students in basic education, 14% master's degree students enrolled in "preschool education and primary school teaching," and 9.3% master's degree students enrolled in "primary school teaching and 2nd grade school teaching in mathematics and experimental sciences".

Stage 2 and stage 3. Validation and comparative sample

Following the initial reliability stage, a sample of 382 participants (mean [M] = 22.60 years, SD = 6.271, age range: 18-54 years) was used to validate the scale and compare the perceptions of self-efficacy in computational thinking between pre-service teachers attending bachelor's and master's degree programs. All participants were from seven universities in mainland Portugal and, as in stage 1, completed the scale during the 2022/2023 academic year. Regarding gender, 93.98% of the participants (359) identified as female, 4.97% (19) as male, and 1.05% (4) chose not to disclose this information. In terms of academic qualifications, 63.4% were pursuing a bachelor's degree in basic education, 6% were enrolled in a master's program in "preschool education," 17% in "preschool education and primary school teaching," 8.6% in "primary school teaching and 2nd grade school teaching in mathematics and experimental sciences," and 5% in "primary school teaching and 2nd grade school teaching in history and geography of Portugal".

Moreover, the characteristics of the participants involved in the studies were recorded using a demographic questionnaire. Data collection for the mentioned study was conducted individually and online, via a link that providing access to the scale on Google Forms.

Instruments

The computational thinking self-efficacy scale (**Appendix A**), developed and validated by Kukul and Karatas (2019), was utilized to assess students' perceptions of self-efficacy in computational thinking. This scale comprises 18 items distributed across four factors: reasoning (items 1, 7, 9, 13, and 16), abstraction (items 2, 5, 10, 15, and 18), decomposition (items 3, 8, 12, and 17), and generalization (items 4, 6, 11, and 14). Each item is rated on a five-point Likert scale, ranging from 1 ("completely disagree") to 5 ("agree fully"). The scale demonstrated excellent reliability, with an overall Cronbach's alpha coefficient of α = 0.884. Reliability scores for individual dimensions were as follows: reasoning (α = 0.772), abstraction (α = 0.774), decomposition (α = 0.701), and generalization (α =0.718), reflecting consistent and robust measurement across all factors.

Ethical Statement

This study adhered to the ethical principles outlined in the Declaration of Helsinki and received approval from the Ethics Committee at the Polytechnic University of Coimbra (approval no. 101_CEIPC/2022, granted on June 24, 2022). Informed consent was obtained from all participants prior to their involvement in the study. Participation was entirely voluntary, and participants were explicitly informed of their right to withdraw at any time without penalty. Confidentiality was rigorously upheld, with all data securely collected and stored in compliance with regulation (EU) 2016/679 of the European Parliament and Council of 27 April (general data protection regulation).

Scale Translation, Adaptation, and Validation Procedures

Stage 1

Translation and adaptation of the scale: Following a thorough analysis of the original version and confirmation of its suitability for this study's context, the process of translating and adapting the scale into Portuguese was initiated. The initial translation was performed by a professional translator to ensure that the meaning of each item was preserved accurately in the Portuguese version. Subsequently, two researchers reviewed the items to refine the language and enhance clarity for the Portuguese audience. A second translator, a native English speaker with no prior exposure to the original scale, then performed a back-translation of the Portuguese text into English. This back-translation was compared with the original text to identify and resolve any discrepancies. Finally, two field experts evaluated both versions to ensure content validity, verifying differences and implementing necessary adjustments (Balbinotti, 2005; Hernandez-Nieto, 2002).

The final wording of the items in Portuguese was defined, confirming the integrity of the instrument, and obtaining the final Portuguese version of the scale. This procedure ensured that the adapted version was faithful to the original instrument and culturally appropriate to the Portuguese context (see **Appendix B**).

Scale stability: The temporal stability (reliability) of the scale was determined by calculating the ICC for the data obtained in the two applications of the test and retest (stage 1). This analysis was conducted by comparing the results of the two applications of the scale: at time 1 (M1) and time 2 (M2), with an interval of 4 weeks between them. The study sample included 43 participants, with an average age of 21.67 years (SD = 5.181), ranging from 18 to 44 years.

Internal consistency: The reliability and internal consistency of the scale were evaluated by calculating Cronbach's alpha for both the overall scale and each individual dimension during both the test and retest phases.

Stage 2

Internal consistency: The analysis of the scale's reliability and internal consistency was conducted using the same methodology employed in stage 1. Cronbach's alpha coefficient was calculated to assess both the overall scale and each dimension individually.

Scale validation: The suitability of the data for exploratory factor analysis was assessed using BET and the KMO index to confirm the adequacy of the correlation matrix (Field, 2018; Marôco, 2021b). Subsequently, factor analysis was conducted using principal component analysis (PCA) with Varimax rotation and Kaiser normalization Factor extraction followed the eigenvalue criterion (greater than 1), complemented by an inspection of the scree plot for a more robust determination of the number of factors. Communalities were analyzed to ensure that the retained items accounted for a significant proportion of the explained variance. Items with factor loadings below 0.4 were removed to enhance the model's quality and interpretability (Field, 2018; Marôco, 2021b).

Additionally, Pearson's linear correlation test was applied to examine the relationships among the identified factors, determining the strength and direction of their associations.

Statistical Analysis

In stage 1, the reliability of the data collected through the application of the scale was evaluated by assessing its internal consistency, measured using Cronbach's alpha coefficient (Pallant, 2020; Pestana & Gageiro, 2014). The alpha coefficients were classified as follows: very good ($\alpha \geq 0.9$); good ($0.8 \leq \alpha < 0.9$); reasonable ($0.7 \leq \alpha < 0.8$); weak ($0.6 \leq \alpha < 0.7$); and unacceptable ($\alpha < 0.6$). To check the temporal stability of the scale, the ICC of the test and retest data was determined. According to the 95% confidence interval for the ICC estimate, values below 0.5 are considered low reliability, moderate reliability between 0.5 and 0.75, good reliability from 0.75 to 0.9 and excellent reliability values above 0.9 (Koo & Li, 2016). To analyze pre-service teachers' perceptions of self-efficacy in computational thinking across the two measurements (test and retest), the Student's t-test for paired samples was conducted after verifying that the necessary assumptions were met (Field, 2018; Marôco, 2021b). The purpose of using this test, which is suitable for contexts where the same samples are analyzed at different times, was to determine whether there are statistically significant differences between the means of two measurements (test and retest). The Kolmogorov-Smirnov test was applied to verify the assumption of normality of the dependent variables. In cases where this assumption was not met, the central limit theorem was used (Marôco, 2021b; Pestana & Gageiro, 2014). The central limit theorem ensures that, for sufficiently large samples ($n \geq 30$), the sample distribution of the mean is close to a normal distribution, allowing the t-test to be used even when the data is not normal. Consequently, the assumption of normality was considered valid (Marôco, 2021b; Pestana & Gageiro, 2014). The effect size for the t-test for paired samples was calculated using Cohen's d, with the following classification (Marôco, 2021b): small effect ($d \leq 0.2$), medium ($0.2 < d \leq 0.5$), large ($0.5 < d \leq 0.8$) and very large ($d > 0.8$).

In stage 2, the construct validity of the scale and the suitability of the data for factor analysis were assessed using the KMO and BET sample adequacy measures (Fávero et al., 2009; Field, 2018; Marôco, 2021b). The KMO test was used to measure the degree of adequacy of the sample, with values greater than 0.70 indicating sufficient adequacy to conduct the factor analysis. In turn, the BET was applied to check whether the correlation matrix between the variables differed significantly from an identity matrix, which would indicate the existence of interdependence between them. In order to proceed with the factor analysis, the BET had to be statistically significant ($p < 0.05$) (Eroğlu, 2008).

The adequacy of the data for exploratory factor analysis was assessed using BET and the KMO index, ensuring the suitability of the correlation matrix. Subsequently, exploratory factor analysis was conducted using PCA with Varimax rotation and Kaiser normalization (Field, 2018; Marôco, 2021b).

The determination of the number of factors followed Kaiser's criterion, retaining only those with eigenvalues greater than 1.0. This was combined with the criterion that each factor should explain at least 5% of the total variance (Field, 2018; Marôco, 2021b). To ensure a more robust definition of the factorial structure, an inspection of the scree plot was also conducted, allowing for the identification of the inflection point in the eigenvalue distribution and the confirmation of the appropriate number of extracted factors.

The adequacy of the extracted factors was assessed based on item communalities, considering a mean communality above 0.6 acceptable for samples exceeding 250 participants, as recommended by Field (2018). Subsequently, items with factor loadings below 0.4 were excluded to ensure that only the most representative items were retained in the factorial structure.

Additionally, the correlation between the extracted factors was analyzed using Pearson's correlation coefficient, allowing for the assessment of the magnitude and direction of their associations. This analysis contributed to a better understanding of the relationships between the underlying constructs and supported the validation of the identified factorial structure. The relationship between the factors was examined using Pearson's linear correlation test (r_p), following the validation of the normality assumption (Marôco, 2021b). The normality assumption was assessed similarly to the t-Student test for paired samples. The classification of relationship strength was based on Hopkins et al. (1996): very weak ($0 \leq r_p < 0.1$), weak ($0.1 \leq r_p < 0.3$), moderate ($0.3 \leq r_p < 0.5$), strong ($0.5 \leq r_p < 0.7$), very strong ($0.7 \leq r_p < 0.9$), almost perfect ($0.9 \leq r_p < 1$), and perfect ($r_p = 1$).

The assessment of model fit quality was conducted through confirmatory factor analysis (CFA), based on the following fit indices: χ^2/df (Chi-squared ratio/degrees of freedom), NFI (normal fit index), CFI (comparative fit index), TLI (Tucker-Lewis index), GFI (Goodness-of-fit index), PCFI (parsimony comparative fit index), PGFI (parsimony goodness of fit index), PNFI (parsimony normal fit index), RMSEA (root mean square error of approximation), SRMR (standardized root mean square residual) (Hair et al., 2019; Hu & Bentler, 1999; Marôco, 2021a). The χ^2/df ratio indicates a very good fit when it is equal to or less than 1, good between 1 and 2, acceptable between 2 and 5, and poor when greater than 5 (Hair et al., 2019; Marôco, 2021a). The values of NFI, CFI, TLI, and GFI indicate a very good fit when equal to or greater than 0.95, good between 0.90 and 0.95, acceptable between 0.80 and 0.90, and poor when below 0.80 (Hair et al., 2005; Kline, 2023; Marôco, 2021a; Marsh et al., 2004). Values of the PCFI, PGFI, and PNFI indices indicate model fit as follows: very good when equal to or greater than 0.80, good between 0.60 and 0.80, and poor when below 0.60 (Marôco, 2021a). The RMSEA value reflects a very good fit when equal to or less than 0.05, good between 0.05 and 0.08, poor between 0.08 and 0.10, and unacceptable when greater than 0.10 (Hair et al., 2005; Hu & Bentler, 1999; Marôco, 2021a). Additionally, an SRMR value below 0.08 indicates a good fit (Marôco, 2021a; Maydeu-Olivares, 2017). Before the CFA was performed, multivariate outliers were assessed using Mahalanobis squared distance (D^2). No observations exhibited values indicative of outliers (p1 and p2 < 0.001). To evaluate the normality of the variables, skewness and kurtosis values were examined, following the criteria ($|Sk| < 3$; $|Ku| < 7$), which confirm normality (Kline, 2023; Marôco, 2021a). In stage 2, the reliability of the data was assessed using Cronbach's alpha as in stage 1.

In stage 3, the comparison of self-efficacy in computational thinking between pre-service teachers enrolled in bachelor's and master's degree programs was analyzed using the Student's t-test for independent samples. This analysis was performed both for the overall scale and for each dimension individually. The assumption of normality was validated before the test was carried out (Marôco, 2021b; Pallant, 2020). This test was used to determine whether statistically significant differences existed in the perceptions of self-efficacy in computational thinking between pre-service teachers enrolled in bachelor's and master's degree programs. The means of the two groups were compared to identify any notable variations. The effect size for the student's t-test for independent samples was calculated using Cohen's *d*, following the same classification used in the Student's t-test for paired samples, as described in stage 1 (Marôco, 2021b).

The analysis of relationship between the dimensions of the self-efficacy scale at each level of education (bachelor's and master's degrees) conducted similarly to the relationship between the factors in stage 2.

All statistical analyses (descriptive statistics, inferential statistics and EFA) were conducted with a significance level of 5% ($p < 0.05$), using the IBM statistical package for the social sciences (version 28, IBM USA). The CFA was performed using the AMOS software (version 28, IBM USA).

RESULTS

Analysis of Stage 1

Internal consistency

To evaluate the reliability and internal consistency of the scale, Cronbach's alpha coefficient was calculated for the total scale and each of its dimensions during both the initial testing stage (M1) and retesting stage (M2). The scale, comprising 18 items, achieved a total Cronbach's alpha (DTotal) of $a = 0.867$ at M1 and $a = 0.877$ at M2, with no improvement observed upon the removal of any

Table 1. Cronbach's alpha coefficients for the scale's dimensions during the testing (M1) and retesting (M2) phases, showing strong reliability across all dimensions and the total scale

	Items	Cronbach's alpha M1	Cronbach's alpha M2
DRacc	1, 7, 9, 13, 16	0.760	0.786
DAbstr	2, 5, 10, 15, 18	0.765	0.771
DDecomp	3, 8, 12, 17	0.803	0.794
DGener	4, 6, 11, 14	0.818	0.809
DTotal	1 to 18	0.867	0.877

Table 2. Descriptive statistics and comparison between total scores at M1 and M2

	M	SD	t	p	d	ES
DTOTAL_M1	69.2791	8.53934	-1.236	0.223	0.188	Small
DTOTAL_M2	70.5581	7.33323				

item. The Cronbach's alpha values for the individual dimension were as follows: reasoning (DRacc) α = 0.760 at M1 and α = 0.786 at M2; abstraction (DAbstr) α = 0.765 at M1 and α = 0.771 at M2; decomposition (DDecomp) α = 0.803 at M1 and α=0.794 at M2; and generalization (DGener) α = 0.818 at M1 and α = 0.809 at M2. These findings indicate consistent reliability across all dimensions and both testing phases (Koo & Li, 2016), as summarized in **Table 1**.

Test-re-test reliability

The temporal stability of the instrument was assessed using the ICC, which measures assessing the consistency of results across the two application points of the scale. The ICC for average measures was 0.778 (F [42, 42] = 4.501; p = 0.001), indicating good reliability according to the criteria of Koo and Li (2016). This indicates that the scale maintains an acceptable level of stability over time.

Furthermore, a paired-samples t-test was conducted to compare the total scores at the two measurement points. The analysis revealed no statistically significant differences between the mean score of DTOTAL_M1 (M = 69.28; SD = 8.54) and DTOTAL_M2 (M = 70.56; SD = 7.33), with t (42) = 1.236 and p = 0.223. The effect size was calculated as d = 0.188, which is considered small, suggesting that any variation between the two measurements is negligible. These findings provide robust evidence of the scale's temporal consistency, as summarized in **Table 2**.

Analysis of Stage 2

Internal consistency of the scale

In line with the methodology employed in stage 1, the reliability and internal consistency of the scale were evaluated by calculating Cronbach's alpha coefficient for both the total scale and its individual dimensions. The scale, consisting of 18 items, achieved a total Cronbach's alpha of α = 0.838, with no improvement observed upon the removal of any item.

Subsequently, Cronbach's alpha was calculated for each of the scale's dimensions, yielding the following results: reasoning (DRacc) α = 0.874, abstraction (DAbstr) α = 0.888, decomposition (DDecomp) α = 0.825, and generalization (DGener) α = 0.816. Similarly, no improvement in alpha was observed when any item was eliminated from the individual dimensions.

These findings demonstrate good internal consistency for the scale, as per the criteria established by Pestana and Gageiro (2014). A comprehensive summary of these results is provided in **Table 3**.

Table 3. Cronbach's alpha coefficients for the total scale and individual dimensions

	Items	Cronbach's alpha
DRacc	1, 7, 9, 13, 16	0.874
DAbstr	2, 5, 10, 15, 18	0.888
DDecomp	3, 8, 12, 17	0.825
DGener	4, 6, 11, 14	0.816
DTotal	1 to 18	0.838

Exploratory factor analysis

The suitability of the data for exploratory factor analysis was initially assessed using the KMO measure of sampling adequacy, which yielded a value of 0.923, indicating a high level of adequacy for factor analysis. Subsequently, BET was performed, producing a statistically significant p-value (p < 0.05), confirming that the correlation matrix was appropriate for factor analysis (**Table 4**).

Table 4. KMO and BET results

Test		Result
KMO		0.923
BET	Approximate Chi-squared	3,005.170
	df	153
	p	0.001

Exploratory factor analysis was conducted using the PCA method, applying Varimax rotation and Kaiser normalization (Field, 2018; Marôco, 2021b). The determination of the number of factors to be extracted followed Kaiser's criterion, which retains only factors with eigenvalues greater than 1.0 and requires that each factor explain at least 5% of the total variance (Field, 2018; Marôco, 2021b). Based on these criteria, four distinct factors were identified, which together explained 60.55% of the total variance. This result indicates that a substantial proportion of the variance in the dataset is accounted for by the factorial structure, supporting the multidimensionality of the scale and suggesting that the identified factors adequately capture the underlying constructs being measured. A detailed analysis of the eigenvalues and the variance explained by each factor is presented in **Table 5**.

Table 5. Total variance explained with 4 factors

Component	Initial eigenvalues			ESSL			RSSL		
	Total	V (%)	CP (%)	Total	V (%)	CP (%)	Total	V (%)	CP (%)
1	7.519	41.773	41.773	7.519	41.773	41.773	4.681	26.005	26.005
2	1.328	7.380	49.154	1.328	7.380	49.154	2.558	14.212	40.217
3	1.118	6.211	55.364	1.118	6.211	55.364	2.471	13.727	53.944
4	0.934	5.187	60.552	0.934	5.187	60.552	1.189	6.608	60.552
5	0.837	4.650	65.202						
6	0.755	4.196	69.398						
7	0.696	3.868	73.266						
8	0.654	3.635	76.901						
9	0.600	3.332	80.233						
10	0.522	2.897	83.130						
11	0.503	2.797	85.927						
12	0.491	2.730	88.657						
13	0.424	2.354	91.010						
14	0.394	2.187	93.198						
15	0.360	1.999	95.197						
16	0.327	1.817	97.014						
17	0.300	1.666	98.680						
18	0.238	1.320	100						

Note. Extraction method: PCA; ESSL: Extraction sums of squared loadings; RSSL: Rotation sums of squared loading; V: Percentage of variance; & CP: Cumulative percentage

Table 6. Component matrix after Varimax rotations

Item	F1	F2	F3	F4	Communalities
1	0.493	**0.632**	-0.129	-0.018	0.659
2	**0.720**	0.333	0.179	0.149	0.684
3	0.329	0.125	**0.613**	0.140	0.519
4	0.388	0.327	-0.053	**0.504**	0.514
5	**0.670**	0.432	-0.004	-0.019	0.636
6	0.393	0.258	0.077	**0.536**	0.514
7	0.441	**0.599**	0.408	-0.134	0.737
8	0.443	0.327	**0.621**	-0.020	0.689
9	0.427	**0.551**	0.234	0.074	0.546
10	**0.627**	0.263	0.314	0.082	0.567
11	0.483	0.245	0.101	**0.512**	0.566
12	0.368	0.129	**0.643**	-0.053	0.568
13	0.385	**0.540**	0.286	0.113	0.534
14	0.472	0.055	0.350	**0.517**	0.616
15	**0.660**	0.189	0.338	-0.076	0.592
16	0.393	**0.589**	0.318	0.116	0.616
17	0.427	0.113	**0.682**	0.102	0.671
18	**0.761**	0.259	0.155	0.002	0.670

Note. Extraction method: PCA; Rotation method: Varimax with Kaiser normalization; & rotation converged in 8 iterations

Subsequently, the scree plot was examined to identify the inflection point in the distribution of eigenvalues, aiding in the validation of the appropriate number of extracted factors. The adequacy of the extracted factors was also assessed based on item communalities. The analysis revealed that nearly all items explained at least 50% of the variance in the original variables, supporting the construct validity of the scale (Field, 2018). For samples larger than 250 participants, a mean communality above 0.6 was considered acceptable, ensuring that the items were well represented by the extracted factors (Field, 2018).

Items with factor loadings below 0.4 were excluded to maintain the robustness of the factorial structure, retaining only the most representative items for the extracted factors (Marôco, 2021b) (**Table 6**).

The distribution of items among the factors was as follows:

Factor 1 (abstraction [DAbstr]): Items 2, 5, 10, 15, and 18

Factor 2 (reasoning [DRacc]): Items 1, 7, 9, 13, and 16

Factor 3 (decomposition [DDecomp]): Items 3, 8, 12, and 17

Factor 4 (generalization [DGener]): Items 4, 6, 11, and 14

This alignment indicates that the factor structure closely mirrors the intended design of the original scale, further supporting its construct validity.

Finally, the correlation between the extracted factors was assessed using Pearson's correlation coefficient, allowing for the analysis of the magnitude and direction of the associations among the factors. This analysis contributed to a better understanding of the relationships between the underlying constructs and supported the validation of the identified factorial structure.

The relationship between the four factors was examined using the Pearson's correlation coefficient. In **Table 7**, it is seen that there is a positive and significant correlation between the four factors.

Table 7. Pearson's linear correlation between the factors

Dimensions	Correlations			
	Abstr	Racc	Decomp	Gener
Abstr		0.791**	0.700**	0.701**
Racc			0.726**	0.718**
Decomp				0.657**
Gener				

Note. $^{**}p < 0.05$

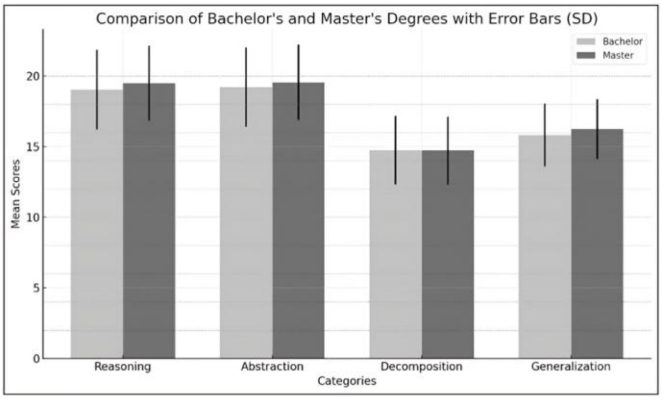

Figure 1. Mean scores with confidence intervals for bachelor's and master's degree groups (Source: Authors' own elaboration)

Confirmatory factor analysis

CFA was performed to evaluate the factor structure and model fit obtained by EFA. The results of the CFA are in line with the EFA analysis, suggesting that the four-factor solution is acceptable (**Table 8**).

Comparison of Pre-Service Teachers' Perception of Self-Efficacy in Computational Thinking

The results of the independent-samples t-test, conducted for both the total scale and each of its dimensions, revealed no statistically significant differences between the bachelor's and master's degree groups.

Figure 1 provides a visual representation of the mean scores for the bachelor's and master's degree groups, along with their respective confidence intervals, facilitating an intuitive comparison.

Additionally, **Table 9** complements this analysis by presenting detailed descriptive statistics and summarizing the comparison between the two groups, further substantiating the lack of significant differences.

The results of Pearson's linear correlation test (r_P) revealed a statistically significant, linear, and very strong positive relationship between the dimensions of the self-efficacy scale for both bachelor's and master's degrees groups.

Table 9. Descriptive statistics and comparison between bachelor's and master's degrees

	Degree	M	SD	t	p	d	ES
DRacc	Bachelor	19.037	2.819	-1.531	0.127	0.163	Small
	Master	19.486	2.654				
DAbstr	Bachelor	19.211	2.807	-1.136	0.257	0.121	Small
	Master	19.543	2.659				
DDecomp	Bachelor	14.744	2.412	0.087	0.930	0.009	Small
	Master	14.721	2.408				
DGener	Bachelor	15.826	2.227	-1.765	0.078	0.187	Small
	Master	16.236	2.107				
DTotal	Bachelor	68.818	9.142	-1.223	0.222	0.130	Small
	Master	69.986	8.710				

Note. ES: Effect size

Table 10. Pearson's linear correlation between the dimensions of the self-efficacy scale at each education level (bachelor's and master's degrees)

Bachelor					
Dimensions	Total	Racc	Abstr	Decomp	Gener
Total		0.912**	0.925**	0.862**	0.852**
Racc			0.800**	0.703**	0.707**
Abstr				0.739**	0.725**
Decomp					0.635**
Gener					
Master					
Dimensions	Total	Racc	Abstr	Decomp	Gener
Total		0.899**	0.904**	0.876**	0.859**
Racc			0.770**	0.702**	0.683**
Abstr				0.706**	0.700**
Decomp					0.705**
Gener					
Total		0.899**	0.904**	0.876**	0.859**

Note. $^{**}p < 0.05$

However, there were notable exceptions: in the bachelor's group, the relationship between the decomposition and generalization dimensions was strong rather than very strong; similarly, in the master's group, the reasoning and generalization dimensions exhibited a strong correlation instead of a very strong one, as detailed in **Table 10**.

Figure 2 provides heatmaps to visualize Pearson's linear correlation coefficients between the dimensions of the self-efficacy scale for the two groups. The intensity of the shading corresponds to the strength of the correlation, with black representing the strongest positive correlations and progressively lighter shades (approaching white) indicating weaker correlations. All shaded areas denote statistically significant results ($p < 0.05$).

The left panel of **Figure 2** illustrates the correlations for the bachelor's degree group, while the right panel depicts those for the master's degree group.

Table 8. Fit indices of the model

χ2(df)	χ^2/df	NFI	CFI	TLI	GFI	PCFI	PGFI	PNFI	RMSEA	SRMR
366.42 (117)	3.132	0.865	0.902	0.872	0.901	0.690	0.688	0.661	0.075	0.0517

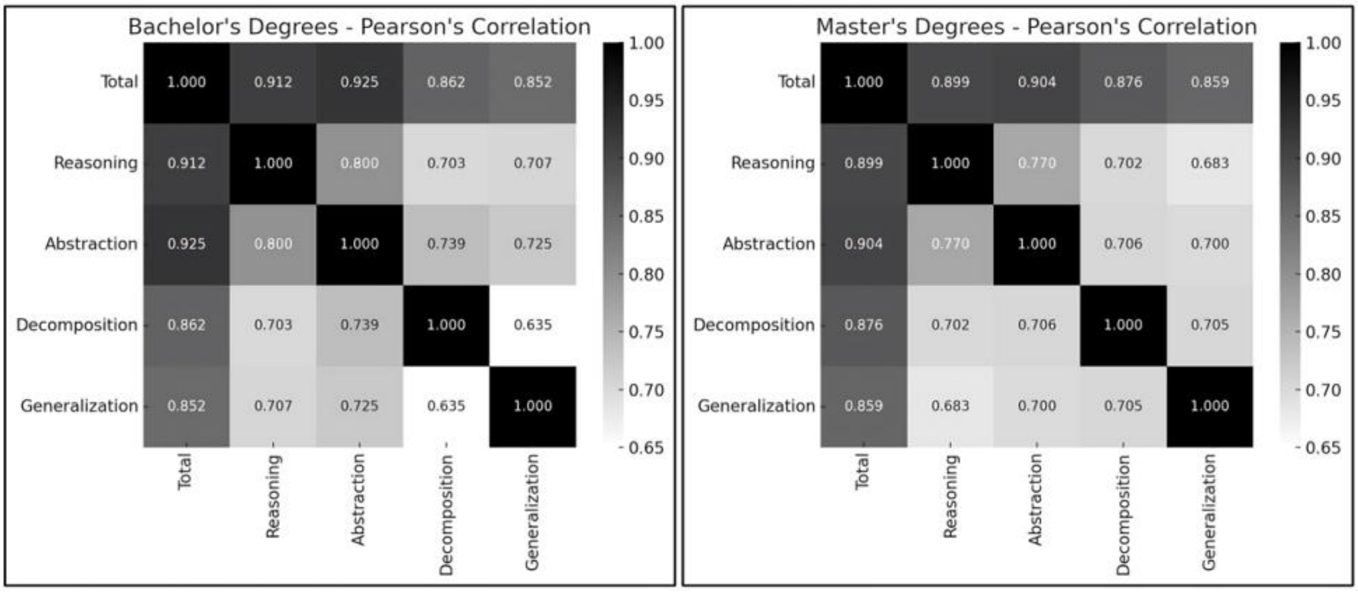

Figure 2. Heatmaps of Pearson's correlation coefficients between self-efficacy dimensions (Source: Authors' own elaboration)

DISCUSSION

The analysis of temporal stability assessed using the ICC (ICC = 0.778; p = 0.001), confirmed that the instrument consistently produced reliable results across different application points. The scale exhibited good internal consistency in both moment 1 and moment 2, as evidenced by the total Cronbach's alpha values of DTotal α = 0.867 and α = 0.877, respectively.

Similarly, Cronbach's alpha values for each dimension demonstrate consistent reliability over time:

- Reasoning (DRacc): α = 0.760 (M1) vs. α = 0.786 (M2)
- Abstraction (DAbstr): α = 0.765 (M1) vs. α = 0.771 (M2)
- Decomposition (DDecomp): α = 0.803 (M1) vs. α = 0.794 (M2)
- Generalization (DGener): α = 0.818 (M1) vs. α = 0.809 (M2)

These findings underscore the instrument's robust reliability and stability across both measurement occasions.

The results of the paired t-test revealed no statistically significant differences between the total scale means at the two application points, highlighting the consistency and stability of participants' perceptions across both instances. This finding is particularly significant for longitudinal studies and repeated measures contexts, where the objective is to track changes in participants' perceptions over time, such as pre- and post-implementation of a training program (Avcı & Deniz, 2022).

Regarding the Portuguese version of the scale, in a sample of 382 individuals, the scale demonstrated acceptable Cronbach's alpha values for both the total scale (α = 0.838) and its individual dimensions (DRacc α = 0.874; DAbstr α = 0.888; DDecomp α = 0.825; DGener α = 0.816). These results align with the criteria established by Pestana and Gageiro (2014) and are like those obtained in the original validation of the scale (Kukul & Karatas, 2019). This level of reliability indicates that the instrument's items are well-correlated, consistently measuring the pre-service teachers' perception of self-efficacy in computational thinking. The validity of the scale was confirmed by the results of the exploratory factor analysis. The KMO value of 0.923 and the BET (p < 0.001) indicated that the data were suitable for factor analysis. The final factor solution resulted in four factors (reasoning, abstraction, decomposition, and generalization) distributed structurally in alignment with the original study regarding dimensionality (Kukul & Karatas, 2019). This structural compatibility with the original scale reinforces the instrument's suitability for measuring the dimensions of the self-efficacy scale–reasoning, abstraction, decomposition, and generalization–in the Portuguese context. Considering the results, it can be concluded that the scale proves to be an appropriate instrument in terms of its reliability and validity, making it suitable for research in various fields. Additionally, the results of the CFA with a four-factor solution revealed a satisfactory model fit, with values for CFI (0.902), GFI (0.901), TLI (0.872), and RMSEA (0.075), within acceptable limits and consistent with the criteria commonly adopted in the literature (Hair et al., 2019; Hu & Bentler, 1999; Marôco, 2021a), corroborating the results of the EFA. The validation of this scale fills a critical gap in research on computational thinking in Portugal, addressing the absence of a reliable instrument to assess participants' self-efficacy perceptions in computational thinking, particularly within the context of initial teacher training. This validation enables the assessment of pre-service teachers' perceptions of their self-efficacy in computational thinking, encompassing those enrolled in both bachelor's and master's programs. This represents a significant contribution to research in this field, as such perceptions can directly impact the effectiveness and success of training program implementations. (Şen, 2023). In relation to the research questions, the analysis revealed no statistically significant differences in self-efficacy perceptions in computational thinking between pre-service teachers enrolled in bachelor's and master's programs. This was true for the total scale (DTotal: t = -1.223, p = 0.222, d = 0.130) as well as for the individual dimensions: (DRacc (t = -1.531, p = 0.127, d = 0.163), DAbstr (t = -1.136, p = 0.257, d = 0.121), DDecomp (t = 0.087, p = 0.930, d = 0.009) and DGener (t = -1.765, p = 0.078, d = 0.187)). The factor analysis for construct validity showed that the scale was in accordance with the theoretical framework and was a valid instrument for assessing the perception of 'computational thinking' self-efficacy. The analysis revealed that there was no statistically significant difference between pre-service teachers enrolled in undergraduate and graduate programs. This supports that the scale works consistently across different academic levels and is an appropriate tool for analyzing self-efficacy perception regardless of the participant's academic degree. Its use in different contexts within initial teacher training will allow for an analysis of perceptions of self-efficacy, enabling comparisons and identification of trends across different groups, as highlighted by Avcı and Deniz (2022).

The lack of statistically significant differences in the perceptions of self-efficacy in computational thinking between pre-service teachers in bachelor's and master's programs may indicate that pre-service teachers, regardless of their stage of training, are similarly underexposed to this concept. This consistency suggests a general absence of structured and explicit instruction in computational thinking across teacher education. The stability in self-efficacy perceptions may indicate that no specific interventions have been implemented to foster or enhance pre-service teachers' self-efficacy in computational thinking, as supported by previous findings (Avcı & Deniz, 2022; Haverback & Parault, 2008). These results align with the study by Avcı and Deniz (2022), which suggests that teacher training programs must aim to develop pre-service teachers' understanding of computational thinking in order to enhance their self-efficacy perceptions in this area. Merely accumulating professional experience is not sufficient to achieve this increase. These findings are also aligned with the study by Graça and Colaço (2024), which highlights the absence of specific training in computational thinking within initial teacher training programs in Portugal. This reinforces the urgent need to develop and implement targeted training programs in this area.

Study Limitations

A notable limitation of this study is the relatively small number of existing studies on the topic, which restricts the current body of knowledge. Specifically, there is a lack of extensive scientific contributions exploring the relationship between self-efficacy and computational thinking abilities, as also observed in Şen's study (2023). This limited research context makes it more difficult to compare the findings of this study with those of similar investigations.

Another potential limitation of the study relates to the composition of the sample, which primarily consisted of female participants. However, given that in Portugal, initial teacher training programs are predominantly attended by women, the sample used in this study is a representative reflection of reality and aligns with the characteristics of the target population, rather than distorting the results.

Lastly, another potential limitation related to the participant group, as the sample was exclusively composed of pre-service teachers enrolled in bachelor's or master's degree programs. Future studies should replicate this research with a more diverse sample, including, for instance, doctoral students and practicing teachers. Such diversification would provide valuable insights into whether differences in perceptions of self-efficacy are influenced by participants' academic levels or professional experiences.

Suggestions for Future Studies

The main objective of this study was to validate a scale to analyze the pre-service teachers' perception of self-efficacy in computational thinking. To achieve this, and in accordance with Nunnally (1978), the sample adhered to the criteria of including at least ten times the number of items on the scale and having more than 300 participants. To strengthen the external validity of the results, it is recommended that future studies replicate this work with a larger and more representative sample of the Portuguese population, facilitating the generalization of the findings to a broader audience.

To advance research and enhance understanding of perceptions of self-efficacy in computational thinking, future studies are encouraged to explore targeted interventions. These interventions should aim to foster the development of computational thinking skills, ultimately seeking to elevate self-efficacy levels in this domain. With this contribution, a valid and suitable scale for initial teacher training in Portugal is now available, allowing for the assessment of the impact of training programs developed for this purpose. This will enable the analysis of the effectiveness of these programs in promoting pre-service teachers' perception of self-efficacy in computational thinking (Avcı & Deniz, 2022).

Furthermore, longitudinal studies could provide deeper insights into how self-efficacy may evolve following the implementation of computational thinking training programs.

Practical Applications

The absence of statistically significant differences in the perception of self-efficacy in computational thinking between pre-service teachers pursuing a bachelor's degree and those pursuing a master's degree highlights a critical gap in research on computational thinking in Portugal: the insufficient teacher training and preparation needed to enable educators to integrate these skills into their teaching practices (Graça & Colaço, 2024; Pinheiro et al., 2023; Ramos et al., 2022). Therefore, it is essential to incorporate training programs focused on developing computational thinking skills into initial teacher education (Bower & Falkner, 2015). For the training to effectively enhance pre-service teachers' self-efficacy in computational thinking, it is essential that it is continuous and long-term, encompassing theoretical, practical, and reflective components (Mason & Rich, 2019; Rodrigues et al., 2024).

CONCLUSIONS

This study translated, adapted, and validated a scale into Portuguese, enabling the assessment of pre-service teachers' perceptions of self-efficacy in computational thinking, thereby addressing the lack of validated instruments for this purpose. The findings revealed no statistically significant differences in the four dimensions of the scale between bachelor's and master's students. Additionally, the relationships among the dimensions followed a consistent trend–significant, linear, and positive–across both levels of initial teacher training.

Given the importance of integrating targeted programs into initial teacher training to enhance self-efficacy in computational thinking, this instrument serves as a crucial tool for monitoring and evaluating such initiatives. Moreover, it provides a solid foundation for future research in Portugal, with additional studies recommended to validate the scale across diverse contexts and samples.

Author contributions: RNR, CC, MA, FM, & SB-C: resources, visualization, and writing–review & editing; **RNR:** conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, software, validation, and writing–original draft; **CC:** conceptualization, data curation, supervision, and validation; **MA:** writing–original draft; **FM:** conceptualization, data curation, funding acquisition, methodology, project administration, software, supervision, and validation; & **SB-C:** conceptualization, formal analysis, methodology, software, supervision, and writing–original draft. All authors agreed with the results and conclusions.

Funding: This work was supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project UIDB/50008/2023 (Instituto de Telecomunicações), UIDB/05198/2023 (Centro de Investigação e Inovação em Educação, inED), UIDB/00194/2023 (CIDTFF), UID/06185/2023 (SPRINT - Sport Physical Activity And Health Research And Innovation Center) and under the doctoral scholarship 2022.09720.BD. Maryam Abbasi thanks the national funding by FCT - Foundation for Science and Technology, PI, through the institutional scientific employment program contract (CEECINST/00077/2021).

Ethical statement: The authors stated that the study was approved by the Ethics Committee at the Polytechnic University of Coimbra on 24 June 2022 (protocol code 101_CEIPC/2022). Written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Ausiku, M., & Matthee, M. (2021). Preparing primary school teachers for teaching computational thinking: A systematic review. In C. Pang (Ed.), *Learning technologies and systems. SETE ICWL 2020 2020. Lecture Notes in Computer Science (), vol 12511* (pp. 202-213). Springer. https://doi.org/10.1007/978-3-030-66906-5_19

Avcı, C., & Deniz, M. N. (2022). Computational thinking: Early childhood teachers' and prospective teachers' preconceptions and self-efficacy. *Education and Information Technologies, 27*(8), 11689-11713. https://doi.org/10.1007/s10639-022-11078-5

Balbinotti, M. A. A. (2005). In order to evaluate the expectations, it is necessary to reflect on the validity of psychological tests. *Aletheia, 1*(21), 43-52.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bower, M., & Falkner, K. (2015). Computational thinking, the notional machine, pre-service teachers, and research opportunities. *Conferences in Research and Practice in Information Technology Series, 160*, 37-46.

Bower, M., Wood, L. N., Lai, J. W., Highfield, K., Veal, J., Howe, C., Lister, R., Mason, R., & Highfield, R. (2017). Improving the computational thinking pedagogical capabilities of school teachers. *Australian Journal of Teacher Education, 42*(3), 53-72. https://doi.org/10.14221/ajte.2017v42n3.4

Butler, D., & Leahy, M. (2021). Developing preservice teachers' understanding of computational thinking: A constructionist approach. *British Journal of Educational Technology, 52*(3), 1060-1077. https://doi.org/10.1111/bjet.13090

Çakir, R., Rosaline, S., & Korkmaz, Ö. (2021). Computational thinking skills of Turkish and Indian teacher candidates: A comparative study. *International Journal of Psychology and Educational*

Studies, 8(1), 24-37. https://doi.org/10.17220/ijpes.2021.8.1.226

Dong, W., Li, Y., Sun, L., & Liu, Y. (2024). Developing pre-service teachers' computational thinking: A systematic literature review. *International Journal of Technology and Design Education, 34*(1), 191-227. https://doi.org/10.1007/s10798-023-09811-3

Durak, H. Y., Uslu, N. A., Bilici, S. C., & Güler, B. (2023). Examining the predictors of TPACK for integrated STEM: Science teaching self-efficacy, computational thinking, and design thinking. *Education and Information Technologies, 28*(7), 7927-7954. https://doi.org/10.1007/s10639-022-11505-7

Eroğlu, A. (2008). Factor analyses. In S. Kalaycı (Ed.), *Multivariable statistic techniques with SPSS applications* (pp. 321-331). Asil Pub.

Esteve-Mon, F. M., Adell-Segura, J., Nebot, M. Á. L. A. L., Novella, G. V., & Aparicio, J. P. (2019). The development of computational thinking in student teachers through an intervention with educational robotics. *Journal of Information Technology Education-Innovations in Practice, 18*, 139-152. https://doi.org/10.28945/4442

Fávero, L. P., Belfiore, P., Silva, F. L., & Chan, B. L. (2009). *Análise de dados: Modelagem multivariada para tomada de decisões* [Data analysis: Multivariate modeling for decision making]. Elsevier.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE.

Gao, X. M., & Hew, K. F. (2022). Toward a 5E-based flipped classroom model for teaching computational thinking in elementary school: Effects on student computational thinking and problem-solving performance. *Journal of Educational Computing Research, 60*(2), 512-543. https://doi.org/10.1177/07356331211037757

Graça, A., & Colaço, S. (2024). Pensamento computacional: Desafios para os professores [Computational thinking: Challenges for teachers]. *Revista UI_IPSantarém, 12*(1), Article e33679. https://doi.org/10.25746/ruiips.v12.i1.33679

Grover, S., & Pea, R. (2013). Computational thinking in K-12: A review of the state of the field. *Educational Researcher, 42*(1), 38-43. https://doi.org/10.3102/0013189X12463051

Hair Jr, J. F., Anderson, R., Tatham, R., & Black, W. (2005). *Análise multivariada de dados* [Multivariate data analysis]. Bookman.

Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Haverback, H. R., & Parault, S. J. (2008). Pre-service reading teacher efficacy and tutoring: A review. *Educational Psychology Review, 20*(3), 237-255. https://doi.org/10.1007/s10648-008-9077-4

Hernandez-Nieto, R. (2002). *Contributions to statistical analysis*. Los Andes University Press.

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

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163. https://doi.org/10.1016/J.JCM.2016.02.012

Kravik, R., Berg, T. K. T., & Siddiq, F. (2022). Teachers' understanding of programming and computational thinking in primary education–A critical need for professional development. *Acta Didactica Norden, 16*(4), Article 23. https://doi.org/10.5617/adno.9194

Kukul, V., & Karatas, S. (2019). Computational thinking self-efficacy scale: Development, validity and reliability. *Informatics in Education, 18*(1), 151-164. https://doi.org/10.15388/infedu.2019.07

Macann, V., & Carvalho, L. (2021). Teachers use of public makerspaces to support students' development of digital technology competencies. *New Zealand Journal of Educational Studies, 56*(SUPPL 1), 125-142. https://doi.org/10.1007/s40841-020-00190-0

Marôco, J. (2021a). *Análise de equações estruturais: Fundamentos teóricos, software & aplicações* [Structural equation analysis: Theoretical foundations, software & applications]. ReportNumber.

Marôco, J. (2021b). *Análise estatística com o SPSS statistics* [Statistical analysis with SPSS statistics]. ReportNumber.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Mason, S. L., & Rich, P. J. (2019). Preparing elementary school teachers to teach computing, coding, and computational thinking. *Contemporary Issues in Technology and Teacher Education, 19*(4), 790-824.

Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(3), 383-394. https://doi.org/10.1080/10705511.2016.1269606

Ministério da Educação. (2021). *Aprendizagens essenciais de matemática* [Essential math learnings]. Ministério da Educação.

Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge. https://doi.org/10.4324/9781003117452

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Pestana, M. H., & Gageiro, J. N. (2014). *Análise de dados para ciências sociais: A complementaridade do SPSS* [Data analysis for social sciences: The complementarity of SPSS]. Edições Sílabo, Lda.

Pewkam, W., & Chamrat, S. (2022). Pre-service teacher training program of STEM-based activities in computing science to develop computational thinking. *Informatics in Education, 21*(2), 311-329. https://doi.org/10.15388/infedu.2022.09

Pinheiro, M. M., Albuquerque, C., Moreira, F. T., Torres, J. V., & Sousa, J. M. (2023). Pensamento computacional na educação: Que sentido faz e que competências promove [Computational thinking in education: What sense does it make and what skills does it promote?] In *Matemática com vida: Diferentes olhares sobre o pensamento computacional* (pp. 9-26). UA Editora. https://doi.org/10.48528/3e5j-1e87

Ramos, J. L., Espadeiro, R. G., & Monginho, R. (2022). *Introdução à programação, robótica e ao pensamento computacional na educação pré-escolar e 1.o ciclo do ensino básico. Necessidades de formação de educadores e professores* [Introduction to programming, robotics and computational thinking in preschool and primary education. Training needs for educators and teachers]. Centro de Investigação em Educação e Psicologia da Universidade de Évora.

Rich, K. M., Spaepen, E., Strickland, C., & Moran, C. (2020). Synergies and differences in mathematical and computational thinking: Implications for integrated instruction. *Interactive Learning Environments, 28*(3), 272-283. https://doi.org/10.1080/10494820.2019.1612445

Rodrigues, R. N., Costa, C., & Martins, F. (2024). Integration of computational thinking in initial teacher training for primary schools: A systematic review. *Frontiers in Education, 9*. https://doi.org/10.3389/feduc.2024.1330065

Román-González, M., Pérez-González, J. C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the computational thinking test. *Computers in Human Behavior, 72*, 678-691. https://doi.org/10.1016/J.CHB.2016.08.047

Şen, Ş. (2023). Relations between preservice teachers' self-efficacy, computational thinking skills and metacognitive self-regulation. *European Journal of Psychology of Education, 38*(3), 1251-1269. https://doi.org/10.1007/S10212-022-00651-8

Sun, L. H., You, X. X., & Zhou, D. H. (2023). Evaluation and development of STEAM teachers' computational thinking skills: Analysis of multiple influential factors. *Education and Information Technologies, 28*(11), 14493-14527. https://doi.org/10.1007/s10639-023-11777-7

Waddington, J. (2023). Self-efficacy. *ELT Journal, 77*(2), 237-240. https://doi.org/10.1093/elt/ccac046

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33-35. https://doi.org/10.1145/1118178.1118215

APPENDIX A: COMPUTATIONAL THINKING SELF-EFFICACY SCALE

Table A1. Computational thinking self-efficacy scale (Kukul & Karatas, 2019)

No	Scale
1	I recognize repetitive structures in data or images.
2	I evaluate the steps necessary for solving the problem from different perspectives.
3	I carry out more than one task at the same time to solve a problem.
4	I distinguish whether a problem I encounter is similar to problems I have encountered before.
5	I analyze the data I collect to solve the problem.
6	I relate problems to real life.
7	I sort data according to their types (text, number, sequence, etc.).
8	I understand whether the problem consists of sub-problems.
9	I have decided whether the data to be used to solve the problem is sufficient.
10	I comment on the data I use to solve the problem.
11	I make connections between the problems I encounter and the problems I have encountered before.
12	If the problem has sub-problems, I manage the solution processes of these sub problems.
13	I find the fastest solution that works correctly among different process steps.
14	I understand how a problem I encounter differs from problems I have encountered before.
15	I organize the data I collect in a way that is more understandable for solving the problem.
16	I decide whether the problem solution I choose is appropriate for the purpose.
17	If the problem has sub-problems, I break it down into smaller sub-problems.
18	I develop different solutions for solving a problem.

APPENDIX B: PORTUGUESE VERSION OF THE COMPUTATIONAL THINKING SELF-EFFICACY SCALE

Table B1. Portuguese version of the computational thinking self-efficacy scale

No	Scale
1	Reconheço estruturas repetidas em dados ou imagens.
2	Avalio os passos necessários para resolver o problema a partir de diferentes perspetivas.
3	Realizo mais do que uma tarefa ao mesmo tempo para resolver um problema.
4	Distingo se um problema que encontrei é semelhante a problemas que encontrei anteriormente.
5	Analiso os dados que recolho para resolver o problema.
6	Relaciono os problemas com a vida real.
7	Classifico os dados de acordo com o seu tipo (texto, número, sequência, etc.).
8	Compreendo se o problema é composto por problemas mais simples.
9	Decido se os dados a utilizar para resolver o problema são suficientes.
10	Descrevo os dados que utilizo para resolver o problema.
11	Estabeleço ligações entre o problema que encontrei e os problemas que encontrei anteriormente.
12	Se o problema contém problemas mais simples, consigo estruturar os processos de resolução desses problemas mais simples.
13	Encontro a solução mais rápida que funciona corretamente entre as diferentes etapas do processo.
14	Compreendo como é que um problema que encontrei difere dos problemas que encontrei anteriormente.
15	Organizo os dados que recolho de uma forma que seja mais compreensível para resolver o problema.
16	Decido se a solução do problema que escolho é adequada ao objetivo.
17	Se o problema tiver problemas mais simples, divido-o em problemas ainda mais simples.
18	Desenvolvo diferentes propostas de solução para resolver um problema.