OPEN ACCESS Freely available online

ISSN 2737-5315
Volume 1, Issue 1
https://www.iikii.com.sg/journal/EIET
Educational Innovations and Emerging Technologies

Article

Measuring Teacher Self-Efficacy for Integrating Computational Thinking in Science (T-SelECTS)

Lautaro Cabrera [1,*], Virginia Byrne [2], Diane Jass Ketelhut [1], Merijke Coenraad [1], Heather Killen [1] and Jandelyn Plane [1]

[1] University of Maryland, College Park, USA
[2] Morgan State University, USA
* Correspondence: cabrera1@terpmail.umd.edu

Received: May 11, 2021; **Accepted:** Oct 15, 2021; **Published:** Nov 20, 2021

Abstract: With computational thinking (CT) emerging as a prominent component of 21st century science education, equipping teachers with the necessary tools to integrate CT into science lessons becomes increasingly important. One of these tools is confidence in their ability to carry out the integration of CT. This confidence is conceptualized as self-efficacy: the belief in one's ability to perform a specific task in a specific context. Self-reported self-efficacy in teaching has shown promise as a measure of future behavior and is linked to teacher performance. Current measures of teacher self-efficacy to integrate CT are limited, however, by narrow conceptualizations of CT, oversight of survey design research, and limited evidence of instrument validity. We designed a valid and reliable measure of Teacher Self-Efficacy for integrating Computational Thinking in Science (T-SelECTS) that fits a single latent factor structure. To demonstrate the instrument's value, we collected data from 58 pre-service teachers who participated in a CT module within their science methods course at a large Mid-Atlantic university. We found evidence of significant development in pre-service teachers' self-efficacy for integrating CT into science instruction. We discuss how the T-SelECTS instrument could be used in teacher education courses and professional development to measure change in self-efficacy.

Keywords: Self-efficacy, Computational thinking, Teacher education, Science teaching

1. Introduction

Computational thinking (CT) is an emerging concept in K-12 education, particularly within science instruction (National Science and Technology Council, 2018). As reflected by current standards (NGSS Lead States, 2013) and recent legislation (e.g., Maryland General Assembly, 2018), educators are increasingly expected to integrate CT into their teaching to prepare students for a computational world. This expectation, however, faces significant challenges. First, the definition of CT is contested (Tedre & Denning, 2016): researchers and educators define CT both as a broad problem-solving strategy that can be applied "everywhere" (Wing, 2006) and as a "range of concepts, applications, tools, and skill sets" (National Research Council, 2010, p. 10). This variety in definitions can make it difficult for teachers to understand the conceptual underpinnings of CT and therefore to clearly bound "what counts" as CT in their classrooms.

A second challenge to integrating CT into formal education is that typical teacher education and professional development programs rarely address concepts related to CT and computing. By 2019, only 19 states required some computer science training for all preservice teachers (Code.org Advocacy Coalition *et al.*, 2019). Therefore, teachers typically have not been prepared to envision CT in their classrooms and effectively integrate it into their instruction. To address these challenges, researchers and teacher educators highlight the importance of developing new ways of equipping teachers with (a) the necessary technological pedagogical content knowledge (Mishra & Koehler, 2006; Mouza *et al.*, 2017; Shulman, 1986) and (b) the confidence to carry out the integration of CT in their teaching (Barr & Stephenson, 2011; Yadav *et al.*, 2014).

Although teacher education efforts with these goals are emerging (e.g., Hestness *et al.*, 2018; Israel *et al.*, 2015; Yadav *et al.*, 2014; Yang *et al.*, 2018), the measurement of these outcomes is still early in its development. From our review of efforts to capture teachers' levels of confidence in integrating CT, current measures are limited by narrow conceptualizations of CT as programming or robotics, scales with insufficient response options in surveys, and double-barreled questions. Thus, in order to measure self-efficacy as an outcome variable, a reliable and rigorously validated measure must be developed. In this paper, we describe the design of a measure of confidence in integrating CT, provide evidence of its reliability and validity, and evaluate the effects of one teacher education intervention on the development of teachers' confidence using this measure. To conceptualize the level of confidence that teachers have in their ability to integrate CT into their instruction, we use the concept of teacher self-efficacy (Bandura, 1977, 1997; Pajares, 1996), which researchers have shown can be an important predictor of teaching performance (Klassen & Tze, 2014) and their students' outcomes (Ashton & Webb, 1986). Therefore, our study focuses on the following research questions:

1. To what extent is the Teacher Self-Efficacy for integrating Computational Thinking in Science (T-SelECTS) scale a valid and reliable instrument?
2. To what extent does pre-service teacher self-efficacy for integrating CT change after participating in a CT module within a science methods course as measured by this instrument?

To situate our research questions within the broader research fields of self-efficacy and teacher education, we review relevant literature below. Particularly, we present the theoretical framework of self-efficacy, explain the importance of self-efficacy for teacher education, and connect these concepts to the emerging challenges of preparing teachers to integrate CT into their instruction.

Self-efficacy refers to a person's belief in her ability to perform specific tasks under given conditions (Bandura, 1977, 1997). As a psychological construct, self-efficacy is highly specific: people have different levels of self-efficacy for different tasks. In other words, capturing a person's self-efficacy around solving mathematical equations may not provide any information on her confidence to teach 1st grade students how to read. Bandura (1993) suggested that people with high self-efficacy in a particular domain can set higher goals for themselves, react more positively to setbacks, and approach difficult tasks within that domain as opportunities for learning.

This link between self-efficacy and how tasks are approached is supported by studies showing that self-efficacy can be a useful predictor of the respondent's performance within the context of the tasks assessed. For example, Milner described how a teacher who had developed a high sense of self-efficacy reacted to critical feedback from her students with persistence and an urge to "step up to the plate" (2002, p. 32). This predictive relationship has also been shown in other educational contexts (Pajares, 1996) and through different measures of teacher self-efficacy and teacher performance (see Tschannen-Moran & Hoy, 2001 for a review). A more recent review of the literature regarding the relationship between self-efficacy, personality, and teacher performance found a significant link between self-efficacy and teaching performance evaluations (Klassen & Tze, 2014).

In the context of science teaching, multiple studies investigated the link between teacher self-efficacy and teaching performance. For example, Haney *et al.* (2002) investigated the relationship between teachers' self-efficacy around science teaching and their ability to effectively teach science as determined by the Horizon Protocol (Horizon Research, 1998). They found that teachers with higher self-efficacy also scored higher on teaching effectiveness. Andersen *et al.* demonstrated that teachers with a high degree of teaching self-efficacy are more likely to implement lessons that are inquiry-based, interdisciplinary, and tied to the real world. On the other hand, low self-efficacy may lead to the implementation of poorly designed, ineffective learning experiences, as seen in the work of Ginns and Watters (1990).

With these studies in mind, it is clear that self-efficacy is an important psychological factor that plays a role in teacher performance. Indeed, models of teacher learning consider teacher beliefs, including beliefs around teaching capability, to be influential in teachers' learning and professional development (Clarke & Hollingsworth, 2002; Gregoire, 2003; Hammerness *et al.*, 2012). Therefore, teacher education and professional development efforts should aim to equip teachers with the necessary confidence in their ability to teach effectively in addition to content knowledge and pedagogical resources. This need may be most essential when teachers are expected to carry out difficult innovations in teaching, such as restructuring science learning to be inquiry-based and centered on argumentation (Duschl & Osborne, 2002), or, most relevant to this study, integrating CT into their science teaching (Barr & Stephenson, 2011).

Considering that self-efficacy is an important construct that should be measured as an outcome of teacher education and professional development efforts, this paper contributes to the field by developing, validating and testing an instrument for measuring teacher self-efficacy around integrating CT into science instruction.

Researchers who argue that all students should learn CT propose K-12 formal classrooms as a natural context to create widespread CT learning opportunities. However, the integration of CT into formal education depends largely on those who create and deliver learning opportunities for students: teachers (Barr & Stephenson, 2011; Guzdial, 2020; Yadav *et al.*, 2014).

Researchers have investigated multiple aspects of the task of preparing teachers to effectively integrate CT into their instruction. Recognizing that current and future teachers already face significant curricular constraints (both at their teacher education programs and at the K-12 level), researchers have attempted to integrate CT into existing coursework and disciplines. For example, multiple researchers have investigated how teachers learn about CT through a technology integration course and develop competencies to infuse CT into their disciplines (Mouza *et al.*, 2017; Pollock *et al.*, 2019) while other studies have focused on understanding how teachers can integrate CT into science education, given the increasing relevance of computational methods in scientific investigations. For instance, researchers have studied how pre- and in-service teachers develop an understanding of computational ideas and CT to integrate them into their science instruction at the elementary (Hestness *et al.*, 2018; Rich & Yadav, 2019; Yadav *et al.*, 2018) middle (Cadieux Boulden *et al.*, 2018), and high school levels (Ahamed *et al.*, 2010). Researchers and teacher educators have also developed frameworks to guide these efforts, identifying the aspects of CT that can be integrated into science (Sengupta *et al.*, 2013; Weintrop *et al.*, 2016). Taken together, these studies and frameworks suggest that, to make effective CT integration a reality, teachers need to develop at least three competencies: (1) an accurate understanding of CT, (2) an understanding of its pedagogical applications within science, and (3) the confidence and agency to carry out its integration into their classroom teaching.

While other work is investigating how to measure teachers' progress in providing teachers with an accurate understanding of CT and its pedagogical applications within science (Mcklin *et al.*, 2019), this study focuses on the third competency: measuring teachers' confidence in integrating CT into their science teaching. When we reviewed previous studies that have attempted to measure educators' level of confidence with CT, we found important limitations in self-efficacy measurement. For instance, a survey originally designed by Yadav *et al.* (2014) aims to measure "comfort with computing" which focuses on teachers' beliefs around their ability to use computers in the classroom and to learn about CT. While these responses may provide some information on teachers' self-efficacy around integrating CT, the survey is not designed to capture the different aspects that can influence that efficacy self-assessment. Particularly, the only two items related to classroom implementation ask teachers to rate their agreement with the fact that CT "can be incorporated in the classroom" (p. 13). Unfortunately, these items do not measure whether teachers feel like *they* can carry out that incorporation—only whether they believe the incorporation is possible.

While multiple studies around CT teacher education have used Yadav's survey (e.g., Leonard *et al.*, 2017; Mouza *et al.*, 2017), other researchers have developed their own instruments. In some cases, these instruments have been limited by design choices contrary to theories of self-efficacy measurement (Bandura, 2006) and survey development (Krosnick & Presser, 2010). For instance, some instruments contained double-barreled questions that forced respondents to pick one answer even if they would answer differently to each *part* of the question. For example, a self-efficacy scale asked participants to determine their agreement to the following statement: "I believe that incorporating computational thinking activities into my teaching could increase student interest in science and technology." (Ahamed *et al.*, 2010). In this case, respondents could have different opinions about whether incorporating CT would increase student interest in science versus technology. This potential ambiguity limits the confidence that each participant interpreted the item similarly.

Another common issue in current measures is the use of insufficient response options to properly capture variability in respondent attitudes. For instance, Yadav's survey analyzed above only contained four response items in a strongly disagree-strongly agree scale—a practice that, albeit common, is not recommended by survey methodologists (Krosnick & Presser, 2010; Schaeffer & Presser, 2003).

Other studies used insufficient items to properly measure the reliability of their instruments. For example, Bower and Falkner (2015) used only one survey item to measure teacher's confidence in "developing their students' CT abilities" (p. 42). Similarly, Curzon *et al.* (2014) used only one item to ask teachers whether they believed the CT workshops they participated in were "confidence building" (p. 92). Studies that did measure reliability of their instruments typically limited these analyses to a Cronbach's alpha level to provide evidence of internal consistency. However, while Cronbach's alpha can show that the survey is capturing a single construct, issues regarding validity (in other words, whether the construct is measuring what it is attempting to measure) remain.

Existing measures of teacher self-efficacy around CT are also limited by narrow conceptualizations of CT. In cases where researchers developed their own instrument, the definition of CT—and therefore the items in their instrument—only fit the specific context of that study. For example, Jaipal-Jamani and Angeli (2017) measured teacher confidence using four items and a 0-100 response scale but their items only asked about the ability to use robotics for classroom instruction. In another study, Bean *et al.* (2015) only created items regarding teachers' confidence in using programming within their classroom teaching. While robotics and programming can certainly be important applications of CT, these instruments are of limited usefulness to researchers who aim to develop more general competencies in integrating CT and need to measure teachers' confidence in their ability to carry out that integration.

In this study, we aim to counter the limitations we identified in the literature and provide a reliable and validated measure of teachers' self-efficacy in integrating CT into science. Below, we describe the context of our study where we developed and tested the instrument, detail the development of the survey, provide evidence of its validity and reliability, and demonstrate its usefulness in measuring self-efficacy within the context of a teacher education CT module.

2. Materials and Methods

The Teacher Self-Efficacy for integrating Computational Thinking in Science (T-SelECTS) scale was developed as part of a National Science Foundation funded research project at a large Mid-Atlantic university that aims to understand how pre- and in-service teachers learn to integrate CT into their elementary science instruction (Hestness *et al.*, 2018; Killen *et al.*, 2020; McGinnis *et al.*, 2019). The study has two main intervention components: (1) a CT module within an existing science methods undergraduate course for pre-service teachers and (2) a professional development (PD) experience where pre-service and in-service teachers work together to develop CT-integrated science lessons for the elementary level. To answer our first research question, we draw on both components of the study. To answer the second research question, we draw exclusively on the CT module and responses from pre-service teachers.

As Fig. 1 depicts, we designed the T-SelECTS instrument and conducted our validation testing in five steps. We present our process and findings in this sequence below.

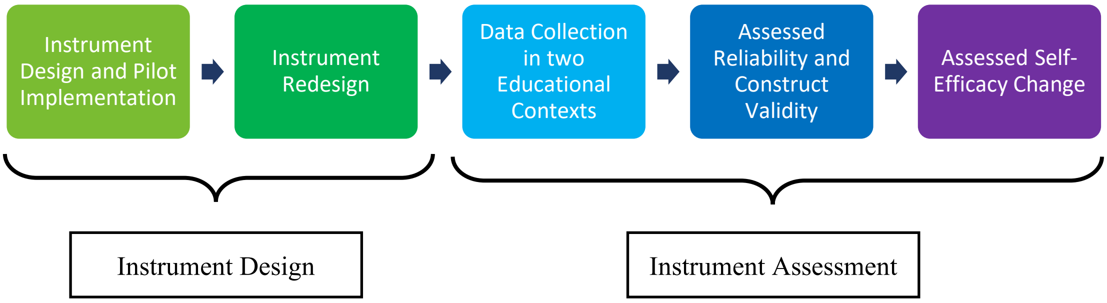

Fig. 1. Instrument Design and Validation Process.

We designed the T-SelECTS scale based on a review of the CT literature that identified what teachers would need to master to effectively integrate CT into their science teaching. Specifically, we referenced the Science Teaching Efficacy Belief Instrument - Form B (STEBI-B; Enochs & Riggs, 1990) to capture the important aspects of science teaching that could be relevant to our measure. This valid and reliable instrument is a widely used standard in science teacher education and has been used to measure teacher self-efficacy in a variety of educational contexts (Deehan, 2017). Based on the STEBI-B, we created new items to attempt to measure teacher perceptions of their technological, pedagogical, and content knowledge around CT and science (Mishra & Koehler, 2006; Mouza *et al.*, 2017; Shulman, 1986). Particularly, we based our instrument on STEBI-B items 5, 8, 12, and 18, which refer to teachers' understanding of and ability to teach disciplinary content and their ability to answer student questions related to that content. We then added items to measure how teachers perceived (a) their own understanding of CT as a concept, (b) their ability to integrate it into their science teaching, (c) their ability to use educational technology to support their teaching, and (d) their ability to engage their students in thinking computationally.

The first version of our survey, which we piloted among 32 pre- and in-service teachers that participated in earlier stages of our project (Cabrera, Jass Ketelhut, *et al.*, 2019; Cabrera, McGinnis, *et al.*, 2019), was a productive initial attempt to measure self-efficacy but displayed many of the limitations that exist in similar instruments. This first version included six items and asked teachers to respond with a four-point Likert-like scale (*Strongly Disagree*, *Disagree*, *Agree*, and *Strongly Agree*). This scale was based on the original STEBI-B instrument, but we removed the middle option (*Uncertain*) to capture teachers' preference towards agreement or disagreement in each item. However, the results of this pilot showed that four response options were insufficient to capture variability in confidence among teachers. Specifically, on three of our items, 75% or more respondents chose the option "Agree" and only a few others chose a different option. In two other items, participants split between two choices, only allowing us to make comparisons between those who "Agreed" and those who "Strongly Agreed"—a differentiation we did not find to be particularly informative.

Our pilot analysis results also revealed other important limitations. For example, through focus groups that we conducted at the end of the course and PD series, we learned that some words in our survey (like "resources") could be interpreted in multiple ways. We also noticed that some items seemed to be uncorrelated with the rest of the survey, indicating that they were irrelevant to self-efficacy development. Therefore, we redesigned the instrument to respond to these limitations and incorporate further recommendations from the literature on self-efficacy measurement.

The first step in redesigning the instrument was to remove any items that seemed to be irrelevant to the measurement of self-efficacy. An initial review of the original instrument indicated that two items asked teachers to rate their ability on tasks that were beyond the expectations of what teachers would need to do to effectively integrate CT into science teaching. One item referred to the ability to find resources around CT, which, in our review, we deemed as irrelevant ability to effectively integrating CT because teachers were *already* being given the necessary resources to perform that task in the two interventions described above—they did not have to find them. The second item regarded the ability to teach peers about CT. While we initially created this item as a task that would be hard for teachers to perform (and therefore provide a range of task difficulty items; Bandura, 2006), we hypothesized that teachers could have high self-efficacy in their ability to integrate CT into their own teaching without necessarily having a related level of self-efficacy in sharing that knowledge with other teachers. Therefore, the two items were removed from the final instrument.

Additionally, we added new items to the instrument that asked about different parts of the process of integrating CT into science instruction. While the piloted instrument only asked teachers whether they could *define* CT, the new survey included items about defining CT, adapting lessons to include CT, creating new lessons that include CT, and answering CT-related student questions. These items covered a wider range of tasks that constitute the larger process of CT integration.

In response to the lack of variability in responses we found using the first version of the survey, we changed the response structure in the final instrument to capture self-efficacy with a continuous measure, following Bandura's (2006) recommendations to create scales with long response continua and multiple items representing tasks of different difficulty. To respond to each item, teachers dragged a slider to mark a number between zero (*Definitely cannot do*) and 100 (*Definitely can do*). The self-efficacy portion of the survey was preceded by four practice items where teachers rated their ability to lift items between 5–300 pounds. These practice items were intended to acquaint teachers with the dragging function of the 0–100 scale in the instrument.

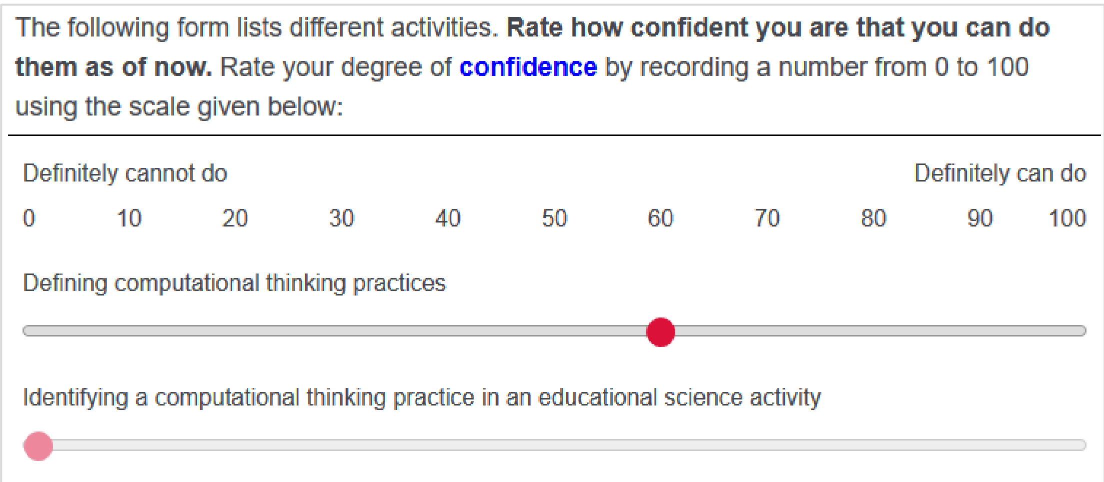

Fig. 2. Survey item scale.

To strengthen the face validity of the scale, we consulted with content experts, asking them to review the instrument and its items. We consulted six faculty members with expertise in elementary science teaching, teacher education, and computer science. These included both senior members of our research team and the study's external advisory board. Their recommendations centered around rewording items to avoid leading questions, deleting items that made assumptions about teachers' backgrounds, and adding "granularity" to the survey by adding items that cover multiple aspects of integrating CT into science instruction. The final instrument included six items, listed in Table 1.

To measure the construct validity, reliability, and usefulness of our instrument, we administered it in two different educational contexts during the Fall 2018 and Winter 2019 semesters: a CT module within a pre-service science methods course and a CT professional development series with pre- and in-service teachers.

All participants were recruited to volunteer in the survey via email and in-class appeals by the research team and were incentivized financially. Participants were required to take the surveys as part of their voluntary participation in the study. They were informed that their responses would be kept confidential and all data would be de-identified for publication or reporting. For pre-service students recruited through the methods course, participation in the study had no impact on students' grades or class performance assessment.

To answer our first research question, regarding the measurement of self-efficacy, we analyzed the pre-test survey responses from all study participants including both in-service and pre-service teachers in both contexts. In-service teachers took the pre-test as part of the beginning of the first session of the PD experience, while pre-service teachers took the pre-test survey at the beginning of the first class of the CT module in their science methods undergraduate course.

To answer our second question, we again surveyed the pre-service teachers at the end of their methods course. We compared their pre- and post-test responses to determine the extent to which pre-service teachers' self-efficacy changed after participating in a science methods course with a CT module.

To answer our first research question, *To what extent is the Teacher Self-Efficacy for integrating Computational Thinking in Science (T-SelECTS) scale a valid and reliable instrument?*, we collected 81 pre-test responses from 18 in-service and 63 pre-service teachers that participated in either the CT module within the undergraduate science methods course or the CT PD series. The majority of our participants identified as white women: 75 women, six men; nine (11.11%) Asian/Asian-American, four (4.94%) Black/African-American, nine (11.11%) Latinx/Hispanic, five (6.17%) Multi-racial, one (1.23%) other, and 53 (65.43%) white. Participants' self-reported teaching experience ranged from having none to 34 years of experience (Mean = 4.79 years, SD = 7.31, Mode = 1 year).

To answer our second research question, we analyzed responses from a subset of participants: those who participated in one component of the study—the CT module in the science methods course. Selecting responses from only one context allowed us to use the instrument to measure change in a *single* intervention. Including participants who participated in two different contexts would make comparisons between timepoints less meaningful. Specifically, we compared the pre- and post-tests of the 58 pre-service teachers who responded to all six items on both surveys: 53 women, five men; seven (12.07%) Asian/Asian-American, three (5.17%) Black/African-American, nine (15.52%) Latinx/Hispanic, one (1.72%) multi-racial, one (1.729%) other, and 37 (63.79%) white. Most (56) of the participants were between 22–26 years old, the remaining two people were 36 and 45.

Because our preliminary normality testing (Shapiro & Wilk, 1965) provided evidence of some nonnormality, we adopted robust statistical methods. To answer our first research question, we conducted an Exploratory Factor Analysis (EFA) using maximum likelihood estimation with robust standard errors (MLR) in MPlus 8.0 to correct for the small sample size, missing data points, and non-normality of our sample (Muthén & Muthén, 1998-2017).

We retained the solution with the best model fit as determined by model fit indices, eigenvalues, scree plot, and factor loadings: a non-significant Chi-square statistic, a standardized root mean square residual (SRMR) value below .08, a comparative fit index (CFI) and Tucker Lewis index (TLI) above .95, and a root mean square error of approximation (RMSEA) score below .06 (Hu & Bentler, 1999; Thompson, 2004). To assess the reliability of the determined scale, we first calculated the internal reliability using Cronbach's alpha, an index of internal consistency (Raykov & Marcoulides, 2011). To produce a robust measurement of reliability, we then calculated the factor replicability, the H-index, seeking a value less than .80 (Hammer, 2016; Hancock & Mueller, 2001).

To answer our second research question, we conducted descriptive analyses and a robust nonparametric version of a paired-samples t-test called the Wilcoxon signed-rank test (Byrne, 2017) in SPSS 25 using the aggregate variables of pre-service teachers' pre-test survey responses and post-test survey responses.

3. Results

3.1 Construct Validity

Before conducting an EFA, we checked for normality, sample size, and sample fitness using SPSS. Four respondents failed to answer all six items (i.e., there were four missing data patterns). A Shapiro-Wilk test of normality provided evidence that some of the items suffered from issues of non-normality ($p < .05$). To determine the appropriateness of the EFA, we calculated the Bartlett test of Sphericity and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA). We found acceptable values for the MSA at .92 and Bartlett test of Sphericity (χ^2 = 691.29, p<.000). From these results, we proceeded to perform an EFA with maximum likelihood estimation with robust standard errors to correct for the missing and nonnormal data in our sample.

We first consulted a scree plot (see Fig. 3) which depicted that a single-factor solution accounted for the majority of the variance in the model. This conclusion was supported by the evidence that only the one-factor solution produced an Eigenvalue greater than one. We concluded that a one-factor solution would be adequate.

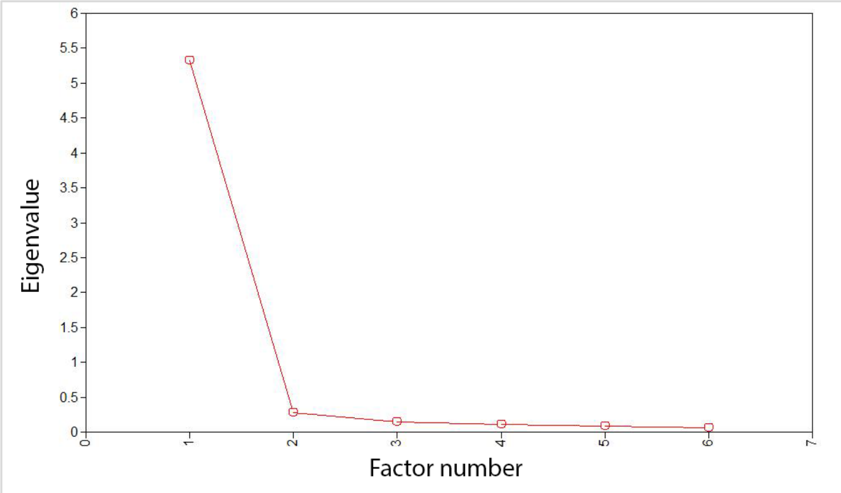

Fig. 3. Scree plot.

We reviewed the model fit statistics, and determined that the one-factor solution demonstrated acceptable fit: the robust Chi-Square statistic (asymptotically equivalent to Yuan and Bentler's T2 statistic; Muthén & Muthén, 1998-2017) was not significant (χ^2 = 16.93, *p*=.05), the SRMR = .02 was well below the threshold of .08 and the CFI = .98 and TLI = .96 were both above .95. The low sample size contributed to the low p-value of the Chi-Square statistic and the high RMSEA = .10 (CI: .00–.18; Kenny *et al.*, 2014). We conclude by retaining the one-factor solution which we interpret as the latent factor of Teacher Self-Efficacy for integrating Computational Thinking in Science (T-SelECTS).

Table 1. Factor Loadings with Oblique Rotation.

Item Number	Item	Factor Loading	Estimated Residual Variances
1	Defining computational thinking practices	0.87*	0.25
2	Identifying a computational thinking practice in an educational science activity	0.94*	0.12
3	Adapting an existing science lesson to include computational thinking	0.95*	0.10
4	Creating an original science lesson that includes computational thinking	0.92*	0.16
5	Engaging students in computational thinking during science instruction	0.96*	0.08
6	Answering student questions regarding computational thinking activities	0.95*	0.10

* significant at 5% level

As presented in Table 1, all six items loaded onto a single factor. All loadings surpassed the .30 cutoff threshold and were significant at the .05 level suggesting that items correlated strongly with the latent T-SelECTS construct. The estimated residual variances are the observed variable variances after accounting for the model variance (Muthén & Muthén, 1998-2017). We calculated the cumulative variance explained to be 93.17% by adding all of the rotated loadings and dividing by the number of items.

3.2. Reliability

The items loading onto the T-SelECTS factor were aggregated into a single variable. The internal reliability of the T-SelECTS factor is acceptable (∝ = .98). The H-value for the T-SelECTS factor was .98; far above the .80 threshold suggesting the items are an acceptable representation of the latent construct (Hancock & Mueller, 2001).

3.3. Using the T-SelECTS to Measure Change in Self-Efficacy

To respond to our second research question, we compared the pre- and post-test responses of pre-service teachers participating in the science methods course where the CT module was integrated. As depicted in Table 2, 58 pre-service teachers completed both the pre and post-test surveys with a post-test mean T-SelECTS score (83.86) nearly double of that the pre-test (37.66). We conducted a Wilcoxon Signed-Rank Test and found a significant increase in pre-service teacher's T-SelECTS scores after participating in the methods course with a CT module, Z= -6.54, p<.001, with a large effect size (r = .61; Cohen, 1988). The median T-SelECTS score increased from 36.50 on the pretest to 87.58 on the post-test.

Table 2. Descriptive Statistics of Pre- and Post-Test T-SelECTS Scores.

Time Point	Mean	Median	SD	SE
Pre-Test (September)	37.66	36.50	23.61	3.10
Post-Test (December)	83.86	87.58	11.80	1.55

4. Discussion

In this study, we present evidence that the T-SelECTS instrument is a reliable and valid measure of teacher's self-efficacy for integrating CT into science teaching. We also demonstrate its usefulness in measuring the impact of a CT module integrated into a science methods teacher education course. As researchers, teacher educators, and policy-makers aim to equip teachers with the necessary tools to integrate CT into science instruction (Barr & Stephenson, 2011; National Science and Technology Council, 2018; Yadav *et al.*, 2014), this survey is a promising instrument to measure teachers' confidence in their ability to carry out that integration. Specifically, our survey addresses limitations in extant work such as insufficient response options, double-barreled questions, and narrow conceptualizations of CT integration. To counter these issues, our instrument uses a 0-100 confidence scale with items that cover a spectrum of tasks of increasing difficulty (Bandura, 2006) associated with the integration of CT into science instruction: from learning about CT as a concept to developing CT-infused science lessons and answering CT-related student questions.

Moreover, the instrument shows promise in its ability to capture variability and change in self-efficacy over time. In our CT module within the teacher education science methods course, we found that, among 58 pre-service teachers who participated in our program and completed our pre- and post-test surveys, there was a significant development in their self-efficacy for integrating CT in future science instruction. Because equipping teachers with the tools to integrate CT into science also involves promoting their confidence to lead that integration, self-efficacy, as measured by the T-SelECTS instrument, could contribute useful information for teacher educators and researchers about the effectiveness of CT integration programs. Therefore, we believe that teacher educators and researchers can use this survey to measure levels of self-efficacy in integrating CT into science in pre- and in-service teacher education contexts.

Future researchers adopting the T-SelECTS instrument could further validate this instrument by increasing the sample size of the participations and randomizing the position of the items in the survey. Additional validation could be aided by analyzing qualitative data that corresponds to survey respondents (Cabrera *et al.*, 2020). Further analysis of the residual correlation between the items is needed to clarify if the similarity of the item phrasing resulted in participants marking similar responses for multiple items.

The self-efficacy gains measured by the T-SelECTS instrument indicate that the CT module was effective in imparting a sense of confidence in teachers in their ability to integrate CT into their future instruction. This increased self-efficacy may lead teachers to set higher goals of CT integration for their lessons and allow teachers to approach the task of CT integration as a learning opportunity to improve their teaching performance (Bandura, 1993). Additionally, it is possible that the increased sense of self-efficacy may lead to better teaching performance and student outcomes (Ashton & Webb, 1986; Klassen & Tze, 2014; Pajares, 1996). However, while our module was successful at increasing teachers' self-efficacy in their ability to integrate CT, the relationship between perceived capability and their implementation of these integrations is beyond the scope of this study. Although studies link self-efficacy with improved teacher performance, we do not address how feelings of competence predict or relate to the ways in which teachers integrate CT into their science teaching in this paper. Future studies could seek to establish a more direct link between pre-service teachers' knowledge, attitudes, and feelings and their classroom practice.

5. Conclusions

While significant attention in being paid to the preparing teachers to integrate CT into their teaching, studies aiming to measure teachers' confidence in their ability to integrate CT into science have important limitations. Thus, there is a gap in the literature on how teacher education and professional development programs should evaluate teacher confidence growth or determine their own program's effectiveness. This study addresses this gap by presenting a valid and reliable measure of Teacher Self-Efficacy for Integrating Computational Thinking in Science (T-SelECTS). Built upon decades of self-efficacy measurement theory, we found the T-SelECTS fits a one-factor model of self-efficacy and has good internal reliability. We contribute the T-SelECTS to the toolset of future researchers and educators to advance our collective goal of supporting CT integration across K-12.

Author Contributions: Conceptualization, Cabrera, Byrne, and Ketelhut; methodology, Cabrera and Byrne; validation, Byrne; formal analysis, Cabrera and Byrne; investigation, Cabrera, Byrne, Ketelhut, Coenraad, Killen, Plane; writing—original draft preparation, Cabrera and Byrne; writing—review and editing, Ketelhut, Coenraad, Killen, and Plane; supervision, Ketelhut.

Funding: This material is based upon work supported by the National Science Foundation under Grant No. 1639891. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A: Complete Final Survey

The following form lists different activities. Rate how confident you are that you can do them as of now. Rate your degree of confidence by recording a number from 0 to 100 using the scale given below:

	0 10 20 30 40 50 60 70 80 90 100
Lifting a 3 pound object ()	
Lifting a 50 pound object ()	
Lifting a 15 pound object ()	
Lifting a 300 pound object ()	

Now that you're familiar with the type of question, let's move on to the next section.

The following form lists different activities. Rate how confident you are that you can do them as of now. Rate your degree of confidence by recording a number from 0 to 100 using the scale given below:

	Definitely cannot do Definitely can do 0 10 20 30 40 50 60 70 80 90 100
Defining computational thinking practices	
Identifying a computational thinking practice in an educational science activity	
Adapting an existing science lesson to include computational thinking	
Creating an original science lesson that includes computational thinking	
Engaging students in computational thinking during science instruction	
Answering student questions regarding computational thinking activities	
Finding resources to facilitate my integration of computational thinking into science teaching	
Teaching colleagues how to include computational thinking into their science teaching	

References

1. Ahamed, S.I., Brylow, D., Ge, R., Madiraju, P., Merrill, S.J., Struble, C.A., & Early, J.P. (2010). Computational Thinking for the Sciences: A Three Day Workshop for High School Science Teachers. *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, 42–46.
2. Andersen, A.M., Dragsted, S., Evans, R.H., & Sørensen, H. (2004). The Relationship Between Changes in Teachers' Self-efficacy Beliefs and the Science Teaching Environment of Danish First-Year Elementary Teachers. *Journal of Science Teacher Education*, *15*(1), 25–38.
3. Ashton, P.T., & Webb, R.B. (1986). *Making a Difference: Teachers' Sense of Efficacy and Student Achievement*. Location: Longman.
4. Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*. https://doi.org/10.1037/0033-295X.84.2.191
5. Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, *28*(2), 117–148. https://doi.org/10.1207/s15326985ep2802_3
6. Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. Location: W. H. Freeman and Company.
7. Bandura, A. (2006). Guide for constructing self-efficacy scales. In *Self-efficacy Beliefs of Adolescents*. Location: Information Age Publishing. (pp. 307–337). https://doi.org/10.1017/CBO9781107415324.004
8. Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: what is Involved and what is the role of the computer science education community? *ACM Inroads*, *2*(1), 48–54. https://doi.org/10.1145/1929887.1929905
9. Bean, N., Weese, J., Feldhausen, R., & Bell, R.S. (2015). Starting from scratch: Developing a pre-service teacher training program in computational thinking. *Proceedings - Frontiers in Education Conference, FIE*, *2014*, 1–8. https://doi.org/10.1109/FIE.2015.7344237
10. Bower, M., & Falkner, K. (2015). Computational Thinking, the Notional Machine, Pre-service Teachers, and Research Opportunities. *Proceedings of the 17th Australian Computing Education Conference*, *January*, 37–46. Available online: http://crpit.com/confpapers/CRPITV160Bower.pdf (accessed on Month Day, Year).
11. Byrne, D. (2017). *Publishing Your Medical Research* (2nd Ed.). Location: Wolters Kluwer.
12. Cabrera, L., Coenraad, M., Mills, K.M., McGinnis, J.R., & Ketelhut, D.J. (2020). Preservice Teachers' Self-Efficacy and Computational Thinking: A Mixed-Methods Approach. *American Educational Research Association Conference 2020 (AERA).*
13. Cabrera, L., Jass Ketelhut, D., Hestness, E.E., Mills, K., & McGinnis, J.R. (2019). Effects of a Computational Thinking Module on Preservice Teachers' Knowledge and Beliefs. *American Educational Research Association Conference (AERA).*
14. Cabrera, L., McGinnis, J.R., Ketelhut, D.J., Hestness, E.E., Mills, K.M., & Jeong, H. (2019). Preservice Teachers' Changes in Self-Efficacy Regarding Computational Thinking. *NARST 92nd Annual International Conference.*
15. Cadieux Boulden, D., Wiebe, E., Akram, B., Aksit, O., Sheridan Buffum, P., Boulden, C., Sheridan, P., & Elizabeth, K. (2018). Computational Thinking Integration into Middle Grades Science Classrooms: Strategies for Meeting the Challenges. *Middle Grades Review*, *4*(3). Available online: https://scholarworks.uvm.edu/mgreview/vol4/iss3/5 (accessed on Month Day, Year).
16. Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(18), 947–967.
17. Code.org Advocacy Coalition, Computer Science Teachers Association, & Expanding Computing Education Pathways. (2019). *2019 State of Computer Science Education Equity and Diversity*. Available online: https://advocacy.code.org/
18. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Location: L. Erlbaum Associates.
19. Curzon, P., Mcowan, P.W., Plant, N., & Meagher, L.R. (2014). Introducing Teachers to Computational Thinking Using Unplugged Storytelling. *WiPSCE '14 Proceedings of the 9th Workshop in Primary and Secondary Computing Education*, 89–92.
20. Deehan, J. (2017). *The Science Teaching Efficacy Belief Instruments (STEBI A and B): A Comprehensive Review of Methods and Findings from 25 Years of Science Education Research* (NV-1 onl). Location: Springer International Publishing. https://doi.org/10.1007/978-3-319-42465-1
21. Duschl, J., & Osborne, J. (2002). Supporting and Promoting Argumentation Discourse in Science Education. *Studies in Science Education*, *38*(1), 39–72. https://doi.org/10.1080/03057260208560187
22. Enochs, L.G., & Riggs, I.M. (1990). Further Development of an Elementary Science Teaching Efficacy Belief Instrument: A Preservice Elementary Scale. *Annual Meeting of the National Association for Research in Science Teaching*. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x
23. Ginns, I.S., & Watters, J.J. (1990). *A Longitudinal Study of Preservice Elementary Teachers' Personal and Science Teaching Efficacy*. 21. http://ezaccess.libraries.psu.edu/login?url=http://search.proquest.com/docview/62566587?accountid=13158
24. Gregoire, M. (2003). Is It a Challenge or a Threat ? of Model Cognition A Dual-Process. *Educational Psychology Review*, *15*(2), 147–179.
25. Guzdial, M. (2020). *Inventing Computing Education to Meet All (Especially Teachers) Undergraduates' Needs: CUE.NEXT Workshops*. BLOG@CACM. Available online: https://cacm.acm.org/blogs/blog-cacm/241939-inventing-computing-education-to-meet-all-especially-teachers-undergraduates-needs-cue-next-workshops/fulltext (accessed on Month Day, Year).

26. Hammer, J. (2016). *Construct Replicability Calculator: A Microsoft Excel-based tool to calculate the Hancock and Mueller (2001) H index*. http://drjosephhammer.com/
27. Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2012). How teachers learn and develop. In *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do.* (Darling-Hammond, L. & Bransford, J., Eds.). Location: Publisher. Vol. 10, pp. 358–389. Available online: https://login.proxy.library.msstate.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2005-13868-010&login.asp&site=ehost-live (accessed on Month Day, Year).
28. Hancock, G.R., & Mueller, R.O. (2001). Rethinking construct reliability within latent variable systems. *Structural Equation Modeling: Present and Future*, 195–216.
29. Haney, J.J., Lumpe, A.T., Czerniak, C.M., & Egan, V. (2002). From Beliefs to Actions: The Beliefs and Actions of Teachers Implementing Change. *Journal of Science Teacher Education*. https://doi.org/10.1023/a:1016565016116
30. Hestness, E., Ketelhut, D.J., McGinnis, J.R., & Plane, J. (2018). Professional Knowledge Building within an Elementary Teacher Professional Development Experience on Computational Thinking in Science Education. *Journal of Technology and Teacher Education*, *26*(3), 411–435.
31. Horizon Research. (1998). *Core Evaluation Manual: Classroom Observation Protocol*. Location: Publisher.
32. Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
33. Israel, M., Pearson, J.N., Tapia, T., Wherfel, Q.M., & Reese, G. (2015). Supporting all learners in school-wide computational thinking: A cross-case qualitative analysis. *Computers and Education*, *82*, 263–279. https://doi.org/10.1016/j.compedu.2014.11.022
34. Jaipal-Jamani, K., & Angeli, C. (2017). Effect of Robotics on Elementary Preservice Teachers' Self-Efficacy, Science Learning, and Computational Thinking. *Journal of Science Education and Technology*, *26*(2), 175–192. https://doi.org/10.1007/s10956-016-9663-z
35. Kenny, D., Kaniskan, B., & Mccoach, D.B. (2014). The Performance of RMSEA in Models With Small Degrees of Freedom. *Sociological Methods & Research*, *44*. https://doi.org/10.1177/0049124114543236
36. Killen, H., Coenraad, M., Byrne, V.L., Cabrera, L., & Ketelhut, D.J. (2020). Reimagining Computational Thinking Professional Development: Benefits of a Community of Practice Model. *Proceedings of the International Conference of the Learning Sciences (ICLS 2020)*, 2125–2132.
37. Klassen, R.M., & Tze, V.M.C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. In *Educational Research Review*. https://doi.org/10.1016/j.edurev.2014.06.001
38. Krosnick, J.A., & Presser, S. (2010). Question and Questionnaire Design. *Handbook of Survey Research*, *94305*, 886. https://doi.org/10.1111/j.1432-1033.1976.tb10115.x
39. Leonard, J., Mitchell, M., Barnes-Johnson, J., Unertl, A., Outka-Hill, J., Robinson, R., & Hester-Croff, C. (2017). Preparing Teachers to Engage Rural Students in Computational Thinking Through Robotics, Game Design, and Culturally Responsive Teaching. *Journal of Teacher Education*, *69*(4), 386–407. https://doi.org/10.1177/0022487117732317
40. McGinnis, J.R., Ketelhut, D.J., Hestness, E.E., Mills, K.M., Jeong, H., & Cabrera, L. (2019). An Examination of Preservice Teachers' Integration of Computational Thinking in Their Elementary School Lesson Plans. *NARST 92nd Annual International Conference.*
41. Mcklin, T., Basu, S., Rutstein, D.W., Weintrop, D., & Yadav, A. (2019). Assessing Computational Thinking: A "Landscape" Symposium About Where We Are at and Where to Go. *American Educational Research Association Conference (AERA)*.
42. Milner, H.R. (2002). *A Case Study of an Experienced English Teacher's Self-Efficacy and Persistence through "Crisis" Situations: Theoretical and Practical Considerations*, *86*(1), 28–35. https://doi.org/10.1037/1076-8998.5.1.84
43. Mishra, P., & Koehler, M.J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
44. Mouza, C., Yang, H., Pan, Y.-C., Yilmaz Ozden, S., & Pollock, L. (2017). Resetting educational technology coursework for pre-service teachers: A computational thinking approach to the development of technological pedagogical content knowledge (TPACK). *Australasian Journal of Educational Technology*, *33*(3), 61–76. https://doi.org/10.14742/ajet.3521
45. Muthén, L.K., & Muthén, B.O. (2017). *Mplus User's Guide* (8th Ed.). Location: Muthén & Muthén.
46. National Research Council. (2010). *Report of a Workshop on the Scope and Nature of Computational Thinking*. https://doi.org/10.17226/12840
47. National Science and Technology Council. (2018). *Charting a Course for Success: America's Strategy for STEM Education* (Issue December).
48. NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States* (Issue November). https://doi.org/10.17226/18290
49. Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*, *66*(4), 543–578. https://doi.org/10.3102/00346543066004543

50. Pollock, L., Mouza, C., Guidry, K.R., & Pusecker, K. (2019). Infusing Computational Thinking Across Disciplines: Reflections & Lessons Learned. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 435–441. https://doi.org/10.1145/3287324.3287469
51. Raykov, T., & Marcoulides, G.A. (2011). Classical Item Analysis Using Latent Variable Modeling: A Note on a Direct Evaluation Procedure. *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 315–324. https://doi.org/10.1080/10705511.2011.557347
52. Rich, K.M., & Yadav, A. (2019). Infusing Computational Thinking Instruction into Elementary Mathematics and Science: Patterns of Teacher Implementation. *2019 Annual Meeting of the Society for Information Technology and Teacher Education (SITE)*.
53. Schaeffer, N.C., & Presser, S. (2003). The Science of Asking Questions. *Annual Review of Sociology*, *29*(1), 65–88. https://doi.org/10.1146/annurev.soc.29.110702.110112
54. Sengupta, P., Kinnebrew, J.S., Basu, S., Biswas, G., & Clark, D. (2013). Integrating computational thinking with K-12 science education using agent-based computation: A theoretical framework. *Education and Information Technologies*, *18*(2), 351–380. https://doi.org/10.1007/s10639-012-9240-x
55. Shulman, L. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, *15*(2), 4–14. http://www.dnv.com
56. Tedre, M., & Denning, P.J. (2016). The long quest for computational thinking. *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, 120–129. https://doi.org/10.1145/2999541.2999542
57. Thompson, B. (2004). *Exploratory and confirmatory Factor Analysis: Understanding Concepts and Applications*. Location: American Psychological Association. https://doi.org/10.1037/10694-000
58. Tschannen-Moran, M., & Hoy, A.W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*. https://doi.org/10.1016/S0742-051X(01)00036-1
59. Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining Computational Thinking for Mathematics and Science Classrooms. *Journal of Science Education and Technology*, *25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5
60. Wing, J.M. (2006). Computational Thinking. *Communications of the ACM*. https://doi.org/10.1145/1227504.1227378
61. Yadav, A., Krist, C., Good, J., & Nadire Caeli, E. (2018). Computational thinking in elementary classrooms: measuring teacher understanding of computational ideas for teaching science. *Computer Science Education*. https://doi.org/10.1080/08993408.2018.1560550
62. Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J.T. (2014). Computational Thinking in Elementary and Secondary Teacher Education. *ACM Transactions on Computing Education*, *14*(1), 1–16. https://doi.org/10.1145/2576872
63. Yang, H., Mouza, C., & Pan, Y.-C. (2018). Examining pre-service teacher knowledge trajectories of computational thinking through a redesigned educational technology course. *Proceedings of International Conference of the Learning Sciences, 1*, 368–375. Available online: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053929497&partnerID=40&md5=1cb2aab021d75653b0aeb22fc88daecc (accessed on Month Day, Year).