MODESTUM
EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(8), em2487
ISSN:1305-8223 (online)
OPEN ACCESS
Research Paper
https://doi.org/10.29333/ejmste/14882

# Adapting the science teaching efficacy beliefs instrument to assess engineering teaching efficacy beliefs of pre-service elementary teachers: Rasch model and confirmatory factor analysis

Ezgi Yesilyurt [1*], Erdogan Kaya [2], Hasan Deniz [3]

[1] Weber State University, Ogden, UT, USA
[2] George Mason University, Fairfax, VA, USA
[3] University of Nevada Las Vegas, Las Vegas, NV, USA

Received 04 April 2024 ▪ Accepted 09 July 2024

**Abstract**

**Background**: The next generation science standards (NGSS) (NGSS Lead States, 2013) place a particular emphasis on the integration of engineering into the science curriculum. Consequently, the NGSS calls on teachers to engage students in engineering practices to facilitate their experience with the engineering design process similar to engineers and develop a more contemporary view of engineering as a discipline. Since engineering education research in K-12 is still in its infancy, there is limited empirical evidence related to how teachers integrate engineering concepts into their classrooms. To assess the quality of teachers' engineering teaching practices, teaching self-efficacy can be used as an indicator of their instructional practices because teaching efficacy beliefs are often associated with greater use of student-oriented instructional practices, increased teacher effort, and other positive instructional behaviors.

**Purpose:** The main aim of this study was to validate an engineering teaching efficacy beliefs instrument (ETEBI) to measure pre-service elementary teachers' engineering teaching efficacy beliefs.

**Design/method**: The science teaching efficacy beliefs instrument version B was modified to develop the ETEBI. The instrument was administered to 561 pre-service elementary teachers. A Rasch model analysis and confirmatory factor analysis (CFA) were conducted on the data obtained from 561 pre-service elementary teachers to provide evidence supporting the validity of the instrument.

**Results:** The Rasch model and CFA suggested a two-factor solution: personal engineering teaching efficacy and engineering teaching outcome expectancy. Also, Cronbach's coefficient alphas for each subscale were measured to assess the internal consistency of the subscales. Based on the analyses, the study provided evidence supporting the reliability and validity of the ETEBI to assess pre-service elementary teachers' engineering teaching efficacy.

**Conclusions:** The ETEBI can be confidently utilized to assess pre-service elementary teachers' engineering teaching efficacy beliefs. It is effective in gauging the current status of their beliefs and/or determining changes in their beliefs as a result of any teacher training and professional development effort.

**Keywords:** pre-service elementary teachers, K-12 engineering, self-efficacy, Rasch model, confirmatory factor analysis

## INTRODUCTION

In the past, science education standards suggested the integration of engineering into K-12 education as a means to improve science learning (e.g., American Association for the Advancement of Science, 1993). However, with the advent of the next generation science standards (NGSS) (NGSS Lead States, 2013), there has been a significant shift toward a more explicit and robust integration of engineering concepts within science education. Engineering and engineering practices have now gained substantial importance and have become an


✉ ezgiyesilyurt@weber.edu **(*Correspondence)** ✉ ekaya3@gmu.edu ✉ hasan.deniz@unlv.edu

**Contribution to the literature**

- This study provides evidence supporting the validity of an instrument, the Engineering Teaching Efficacy Beliefs Instrument (ETEBI), addressing a significant gap in measurement tools for assessing pre-service elementary teachers' engineering teaching self-efficacy.
- The comprehensive validation process using both Rasch model analysis and confirmatory factor analysis (CFA) as well as Cronbach's alpha establishes the reliability and validity of the ETEBI, making it a robust tool for future research.
- The ETEBI offers significant benefits to researchers, teacher educators, and policymakers by providing a reliable tool for evaluating the impact of interventions or programs in engineering education and for supporting the development and improvement of engineering education programs by effectively assessing pre-service elementary teachers' engineering teaching efficacy.

essential component of K-12 science education. One of the core recommendations of the NGSS is the integration of the engineering element into science instruction, not only as a pedagogical strategy that offers a meaningful context for science learning but also as a fundamental concept in and of itself.

However, the integration of engineering in K-12 presents challenges, especially since the NGSS requires K-12 teachers to adjust their curricula to engage their students in engineering practices and to introduce it as a distinct disciplinary core idea. Even though elementary teachers are expected to introduce engineering to their students from elementary grades onward, they are not adequately prepared due to several reasons including lack of training, background knowledge, resources, etc. (Pleasants & Olson, 2019; Purzer et al., 2022; Trygstad et al., 2020).

Considering these challenges, it is crucial to take into account teaching self-efficacy, which is an essential construct for identifying and predicting teachers' instructional practices. Teaching self-efficacy is considered a teacher's belief in their ability to effectively teach a specific subject or skill. It plays a crucial role in shaping educators' attitudes and behaviors in the classroom, ultimately influencing the quality of instruction provided to students. In the context of engineering education, understanding the teaching self-efficacy of elementary teachers is of paramount importance, as it directly impacts their ability to successfully integrate engineering concepts into their teaching practices. Building on this premise, this study aimed to develop and establish a valid and reliable engineering teaching efficacy belief instrument (ETEBI) specifically tailored for pre-service elementary teachers. The ETEBI assesses the extent to which future educators feel confident in their ability to teach engineering-related concepts effectively. Reflecting on the broader implications of this work, focusing on this unique demographic addresses a crucial gap in existing educational measures and aligns with the emphasis that the NGSS places on engineering in K-12 education, particularly in elementary teacher education.

## THEORETICAL FRAMEWORK

Self-efficacy is broadly defined as the perceived capability to successfully act. This construct is grounded in Bandura's (1977) social cognitive theory which posits that self-efficacy beliefs motivate individuals to take particular actions necessary to achieve a goal and that this construct could be used as a variable to make predictions about one's future behavior (Bandura, 1977). This belief can influence how individuals approach challenges and whether they persist amidst challenges. Individuals who have high self-efficacy are more inclined to take on difficult tasks and persist when confronted with obstacles, whereas those who have low self-efficacy often avoid difficult situations and are prone to disengage. Self-efficacy theory has been extensively studied and has been found to predict various outcomes, including academic achievement, job performance, and health-related behaviors.

Bandura (1977) proposed that behavior is influenced by two factors: outcome expectancy and self-efficacy expectancy. Bandura (1977) identified efficacy expectancy as the belief in an individual's capability to attain a targeted outcome, and outcome expectancy as the belief about the consequences that will result from an individual's behavior. Bandura (1977) argued that self-efficacy expectancy and outcome expectancy are cognitive factors influencing human motivation and action. Bandura (1977) suggested that efficacy expectancy and outcome expectancy be considered as two separate yet related constructs. Specifically, in tandem with an increase in teachers' personal efficacy in providing effective instruction on a specific subject, so too is their belief in the potential beneficial outcomes achieved through these effective instructions. However, this is not always the case. For example, within the context of teacher education, teachers might believe that effective pedagogical practices significantly contribute to student achievement, which reflects a high outcome expectancy. But at the same time, they might doubt their own ability to implement these practices due to perceived gaps in their pedagogical knowledge and skills, indicating low personal self-efficacy. Conversely, teachers may feel confident in their pedagogical abilities (high personal self-efficacy) while simultaneously questioning the overall impact of these practices on student achievement (low outcome expectancy).

Bandura's (1977) work on self-efficacy beliefs has been foundational to the research on teacher self-efficacy. Teacher self-efficacy is considered as teachers' beliefs in their abilities to influence student outcomes and is a major predictor of teachers' job satisfaction, commitment, and decision-making (Tschannen-Moran et al., 1998). Tschannen-Moran et al. (1998) described teaching self-efficacy as a "teacher's belief in his or her own capability to organize and execute the courses of action required to successfully accomplish a specific teaching task in a particular context" (p. 233). Tschannen-Moran et al. (1998) introduced a theoretical model explaining how teaching self-efficacy is formed through the interaction between an assessment of personal teaching capabilities and an evaluation of teaching practices within a specific educational context. The resulting self-efficacy beliefs affect teachers' instructional decisions and classroom practices. It has long been established that self-efficacy is associated with teachers' effectiveness, motivation (e.g., Albion & Spence, 2013; Calkins et al., 2024; Demir, 2020; Karpudewan et al., 2023; Klassen & Tze, 2014; Sehgal et al., 2017), effort, enthusiasm (e.g., Burić & Kim, 2020; Huang et al., 2019; Lazarides et al., 2018; Ma et al., 2022), and student outcomes in the classroom (e.g., Ashton & Webb, 1986; Bandura, 1997; Perera & John, 2020; Shahzad & Naureen, 2017). For instance, teachers who perceive themselves as more efficacious in their ability to teach a specific subject tend to engage in effective teaching practices (e.g., practices aligned with constructivist learning approaches) (Boz & Cetin-Dindar, 2021; Livers et al., 2020; Temiz & Topcu, 2013). Research indicates that more self-efficacious teachers are more receptive to new reforms and curriculum changes, more inclined to adopt new teaching methods, and more experimental in catering to student needs (e.g., Livers et al., 2020; Nie et al., 2013; Tschannen-Moran et al., 1998). Nie et al. (2013), for instance, showed a significant positive association between self-efficacy beliefs and reported constructivist instructional practices by examining the beliefs and reported practices of 2,139 in-service primary school teachers. Similarly, Lucero et al. (2013) found in their classroom observation study that teachers who had high self-efficacy were more inclined to encourage open inquiry over guided inquiry.

Moreover, the impact of teacher self-efficacy extends to student outcomes as well. Several scholars have provided evidence of the relationship between teachers' self-efficacy beliefs and students' outcomes (e.g., Tschannen-Moran & Hoy, 2001; van Uden et al., 2014; Zee & Koomen, 2016). van Uden et al. (2014), for example, indicated that teacher self-efficacy influences students' engagement in multiple dimensions: behavioral, emotional, and cognitive. Additionally, several studies have shown that student performance increased in classrooms with teachers who were more confident about their teaching effectiveness (e.g., Midgley et al. 1989; Mojavezi & Tamiz, 2012; Zee & Koomen, 2016). This can be attributed to the fact that highly efficacious teachers offer more opportunities for student participation in the learning process.

## LITERATURE REVIEW

### Teaching Efficacy Surveys

The relationship between teacher self-efficacy and teacher behavior and effectiveness is strongly anchored in the literature (e.g., Bandura, 1997; Karpudewan et al., 2023; Klassen & Tze, 2014; Tschannen-Moran & Hoy, 2001). In this regard, several self-efficacy scales have been developed using Bandura's (1977) social cognitive theory for assessing teachers' beliefs about teaching efficacy over the past two decades.

The historical attempts to develop a teaching efficacy scale focused on improving the validity and reliability of the existing scales and determining the meaning of the components of the scales. These scales were aligned with the different conceptualizations of teaching efficacy: Rotter's (1966) social learning theory and Bandura's (1977) social cognitive theory. Rotter's (1966) instrument was developed based on the attribution-based theory of locus of control which is broadly defined as one's beliefs about the extent to which the outcome of events (failures or successes) is dependent on individuals' own behavior (Rotter, 1966). Rotter (1966) presumed that individuals may vary in their beliefs regarding whether the outcomes are determined by external control or as a consequence of their own actions. The earlier attempts to develop the teacher efficacy scale (TES) made by research and development corporation (RAND) researchers followed Rotter's (1966) locus of control theory as a conceptual basis. Building on that theory, RAND researchers suggested that teachers who have confidence in their ability to have an effect on student learning believe that the outcomes of their teaching are in teachers' hands or internally controlled. Conversely, teachers who believe external factors have a greater influence may feel they lack control over student learning outcomes (Tschannen-Moran et al., 1998). Two items were added to an existing survey aiming to measure teachers' perceptions about their own teaching capabilities by the RAND researchers:

> **Item 1:** "When it comes right down to it, a teacher really can't do much because most of a student's motivation and performance depends on his or her home environment."
>
> **Item 2:** "If I really try hard, I can get through to even the most difficult or unmotivated students."

This ad-hoc addition improved the survey, which was designated to predict teachers' performances. Gibson and Dembo (1984) pointed out that these items are aligned with Bandura's (1977) personal and outcome expectancy conceptualizations. To improve the reliability issues associated with this two-item scale, three TESs which are grounded in Rotter's (1966) theory were developed:

(1) Guskey's (1981) 30-item responsibility for student achievement scale assessing teachers' assumptions about their responsibility for their students' success and failure,

(2) Rose and Medway's (1981) 28-item teacher locus of control scale measuring teachers' perceived responsibility for students' achievements, and

(3) Ashton et al.'s (1982) scale addressing social desirability.

These scales did not gain wide acceptance among scholars (Tschannen-Moran & Hoy, 2001). Drawing on Rotter's (1966) theory, Bandura (1977) posited that an individual's behavior is shaped by generalized expectancies for control along with beliefs in their own abilities, or self-efficacy, to carry out those behaviors in specific situations. As a result, it is essential to consider teachers' personal beliefs about their teaching capabilities, in addition to the students' outcomes, when developing teacher efficacy instruments. Bandura (1977) further suggested that even though individuals may be aware that certain practices bring about targeted outcomes (outcome expectancy), this knowledge becomes nearly inefficacious if they do not believe they have the capabilities to perform such actions (personal self-efficacy). This means that personal self-efficacy is the most essential element of individuals' behavior, which eventually helps predict their outcome expectancy.

After Bandura's (1977) self-efficacy construct gained prominence, several researchers attempted to reconcile Rotter's (1966) theory and Bandura's (1977) construct of self-efficacy to develop instruments measuring teacher efficacy (e.g., Ashton et al., 1984; Gibson & Dembo, 1984). In this line of research, Ashton et al. (1984) developed context-specific vignettes, illustrating situations that teachers could face. Teachers are asked to make judgments about their performance in managing particular situations. However, this scale was not widely accepted in education research. Gibson and Dembo (1984) also attempted to propose a 30-item TES based on Bandura's (1977) efficacy constructs. The factor analysis resulted in a two-factor solution, which Gibson and Dembo (1984) argued aligned with Bandura's (1977) expectancies. Consequently, they labeled these two constructs as personal teaching efficacy and general teaching efficacy. While studies using the TES confirmed its two-factor solution for teacher efficacy, continued research has revealed several inconsistencies within the TES items (Coladarci & Fink, 1995; Guskey & Passaro, 1994). For example, Guskey and Passaro (1994) argued that the dimensions of Gibson and Dembo's (1984) instrument were more closely associated with the locus of control theory than the self-efficacy theory. Therefore, there were several attempts to modify Gibson and Dembo's (1984) scale. For example, Guskey and Passaro (1994) modified Gibson and Dembo's (1984) scale, labeling efficacy beliefs concerning the impact of teachers on student learning as "internal" and factors outside teachers' control as "external".

At that time, consensus on the conceptualization of the teaching efficacy constructs was lacking. As Henson (2001) pointed out, the studies on teacher efficacy were going through "an adolescent identity crisis" (p. 10). Ashton et al. (1982), for instance, proposed two dimensions for teachers' efficacy: teaching efficacy, which pertains to teachers' expectations of student outcomes related to their teaching, and personal teaching efficacy, reflecting teachers' judgments about their abilities to execute teaching-related tasks and activities. Ashton and Webb (1986) argued that personal teaching efficacy and outcome expectancy could function independently. Namely, while some teachers might endorse the view that teaching can have an impact on student learning even when they themselves lack the necessary skills to make the desired impact, others might believe that teaching, in general, does not have a strong effect on student learning, but consider their own teaching as an exception. On the other hand, Soodak and Podell (1993) claimed that teaching efficacy encompasses three distinct dimensions: teaching efficacy, personal efficacy, and outcome efficacy. While teachers' personal efficacy refers to their belief in possessing the necessary teaching skills, outcome efficacy concerns the achievement of desirable student outcomes as a result of these skills, and teaching efficacy is the teacher's ability to override external influences. Tschannen-Moran and Hoy (2001) addressed the issues associated with existing instruments for measuring teacher efficacy beliefs and developed a more reliable teachers' sense of efficacy scale.

Contrary to Bandura's (1977) theory, teacher efficacy was regarded as a general belief about teaching rather than a domain-specific construct, leading to the initial measurements not being specifically designed for particular teaching domains and contexts. On the other hand, general measurements developed to measure teachers' efficacy have been shown to possess a low predictive discriminant validity (e.g., Coladarci & Fink, 1995; Pajares, 1996), as teacher efficacy varies and depends on specific situations and content areas (Ross et al., 1999). Therefore, as Bandura (1977) argued, efficacy judgments are dependent on the context. In line with this thinking, building on the work of Gibson and Dembo (1984), Enochs and Riggs (1990) developed a context-specific instrument, the science teaching efficacy belief instrument version B (STEBI-B), to examine the efficacy of teaching science. In their study, the data collected from 212 pre-service elementary teachers indicated two subscales: The personal science teaching efficacy (PSTE) scale and the science teaching outcome expectancy (STOE) scale. The PSTE scale consists of 13 items measuring teachers' perceptions of their ability to teach science effectively, explain science concepts, and integrate scientific experiments into their classrooms. The STOE includes 10 items measuring teachers' perceptions of how their science teaching impacts students' achievements. The Cronbach's alpha coefficient of internal consistency for the PSTE scale was found to be .90 and for the STOE scale was found to be .76.

Using the STEBI-B as a basis, several content-specific efficacy instruments were created. To measure chemistry teaching efficacy beliefs, for instance, the STEBI-CHEM was designed by Rubeck and Enochs (1991). In the context of mathematics teacher education, the mathematics teaching efficacy belief instrument (MTEBI) was created to examine pre-service teachers' math teaching efficacy beliefs (Enochs et al., 2000). To assess computational thinking, authors adopted and modified the STEBI instrument (Kaya et al., 2019, 2020). Last but not least, Ritter et al. (2001) created an instrument to measure teachers' efficacy beliefs regarding equitable science teaching, specifically aimed at evaluating teachers' effectiveness in teaching science to underrepresented student groups.

In the context of engineering education, Yoon et al. (2014) highlighted the absence of an instrument to examine practicing (in-service) teachers' engineering teaching self-efficacy in K-12 settings and developed a scale to measure this construct. They adopted several teacher self-efficacy instruments to develop the teaching engineering self-efficacy scale. Yoon et al. (2014) conducted exploratory and confirmatory factor analyses based on the data from 434 K-12 teachers. The analyses yielded four factors including

(a) engineering pedagogical content knowledge,

(b) engineering engagement,

(c) engineering disciplinary self-efficacy, and

(d) outcome expectancy (p. 478).

They achieved a high internal consistency for each factor. On the other hand, in their study, Yoon et al. (2014) pointed out that the consistency of the scale's validity across different grade levels was not determined. While this study represents the first step towards developing a reliable and valid instrument for the engineering teaching efficacy construct, the instrument was specifically developed to assess the efficacy beliefs of K-12 in-service teachers. In that sense, the need for this study arises from the fact that there is no, to the best of the authors' knowledge, an instrument measuring specifically pre-service elementary teachers' engineering teaching efficacy beliefs. In addition, Bandura (1977) asserted that self-efficacy is a situation-, domain- and task-specific construct. In that regard, given that elementary teaching has its own unique context, different from middle and high school teaching contexts (e.g., Hammack & Ivey, 2017b; Savran & Çakıroğlu, 2003), elementary teachers' efficacy information should be assessed by a grade and content-specific instrument.

From this perspective, following the work of Enochs and Riggs (1990), this study attempted to develop and provide supporting evidence for the validity of an engineering teaching self-efficacy instrument specifically designed for pre-service elementary teachers. In recent applications, some researchers have adapted the science teaching efficacy belief instrument (STEBI) by changing the context from science to engineering (e.g., Kaya et al., 2019; Yesilyurt et al., 2021), thereby laying the exploratory groundwork for the development of the ETEBI. This modification draws inspiration from and relies on the previously established validity of the STEBI across various disciplines, such as chemistry and math as discussed above. Notably, the studies employing the adapted STEBI have yielded comparable factor loadings, indicating cross-disciplinary validity, and providing a foundation for ETEBI. While previous studies using the ETEBI have presumed evidence for validity based on its established use in other disciplines, the researchers emphasized the need for rigorous validation specific to the engineering context. The present study aims to address this gap by providing a robust validation process to offer evidence for the validation of the ETEBI instrument, designed for pre-service elementary teachers. The psychometric properties of this instrument were assessed through Rasch analysis and confirmatory factor analysis (CFA), providing insights into its reliability and validity. Given that the STEBI-B developed by Enochs and Riggs (1990) serves as the most widely utilized tool in the field of pre-service elementary science education, the present study used the STEBI version B as a basis for designing the engineering-specific teaching efficacy beliefs instrument for pre-service elementary teachers.

## The Place of Engineering in Pre-College Education

The NGSS places special emphasis on engineering, treating it both as an integrated component and a distinct subject within K-12 science education. The NGSS suggests integrating engineering as a pedagogical technique for teaching science, potentially providing valuable opportunities for students to apply their scientific knowledge to practical problems. The NGSS also presents engineering design as an essential, stand-alone subject. Alongside engineering practices, the standards identify core ideas of *engineering design* as a specific domain of knowledge pertaining to engineering practices. Specifically, the standards set forth 36 performance expectations focused on engineering, with 14 explicitly connected to the core idea of *engineering design*. Therefore, engaging in engineering practices and exploring engineering core ideas afford students the chance to understand the epistemic aspects of engineering. The standards require learners not only to engage in engineering design processes but also to learn the nature of engineering design.

This new vision of science education may bring new challenges for teachers, given that teachers do not have formal training in engineering and lack pedagogical knowledge about how to teach it (e.g., Hynes et al., 2017; National Academies of Sciences, Engineering, and Medicine [NASEM], 2020). The 2012 national survey of science and mathematics education (NSSME) revealed that fewer than 5% of elementary science teachers had exposure to engineering college coursework and that nearly 80% of them reported that they were insufficiently prepared to teach engineering (Trygstad et al., 2013). The recent results of the 2018 NSSME are not different from those of the 2012 NSSME. The national survey revealed that 51% of elementary school teachers reported not feeling adequately prepared to integrate engineering into their teaching curriculum, while only 3% felt well prepared to do so (Banilower et al., 2018).

Studies attempting to examine elementary teachers' perceptions of engineering revealed that teachers had various misconceptions about engineering and lacked pedagogical content knowledge (e.g., Hammack & Ivey, 2017a; Hsu et al., 2011; Kuvac & Koc, 2023). Kuvac and Koc's (2023) study on pre-service teachers, for example, illustrated that the teachers held inadequate knowledge about engineering work and engineering design. Specifically, they often associated engineering primarily with manual tasks, such as constructing and operating machinery or vehicles before these systems are put into practice. In another study, Hammack and Ivey (2017a) indicated that while many teachers described engineers primarily as designers or creators, they frequently emphasized physical tasks such as building and fixing machines. Additionally, there was a significant amount of uncertainty and limited understanding among teachers regarding what engineers do and the engineering design process, with some confusing it with the scientific method.

Given the insufficient conceptions of engineering that K-12 teachers hold, several scholars measured engineering teaching efficacy beliefs to determine the impact of limited engineering knowledge and prior training on teachers' confidence in teaching engineering. For example, research by Hammack and Ivey (2017b), investigated 542 elementary teachers' engineering/engineering design knowledge and engineering/engineering design teaching efficacy. The study revealed minimal exposure to engineering in the teachers' backgrounds, including in engineering and engineering design coursework or professional development programs. In parallel, the analysis showed that teachers possessed low self-efficacy beliefs concerning their knowledge of engineering and engineering design, as well as their capability to use their engineering knowledge in teaching and to influence students' learning. Several scholars also explored pre-service elementary teachers' perceptions and engineering teaching efficacy and indicated that pre-service elementary teachers not only held a naive view of engineering as a discipline but also had low engineering teaching efficacy beliefs (Kaya et al., 2017; Yesilyurt et al., 2021). In a similar vein, Kang et al. (2018) investigated elementary teachers' conceptions of the NGSS science and engineering practices and found that although teachers reported using engineering design activities, none was able to articulate the connection between their engineering design activities and the NGSS engineering practices.

Building on the insights from these studies, which highlight in-service and pre-service elementary teachers' misconceptions about engineering and their low self-efficacy in teaching engineering concepts, this research aims to develop a valid and reliable instrument specifically designed to provide a deeper understanding of the status of pre-service elementary teachers' self-efficacy in teaching engineering. This instrument will contribute to targeted interventions that can enhance their efficacy by indicating specific areas that need improvement.

## METHODOLOGY

### Participants

Participants were chosen using a combination of purposive and convenience sampling methods. Given that the study aimed to develop an instrument for pre-service elementary teachers, the study employed purposive sampling to ensure the sample accurately reflected this group. In this sense, the study involved 561 pre-service elementary teachers. Additionally, convenience sampling was used to select participants based on their accessibility and availability. The participants were selected from an R1 university, classified by the Carnegie classification of institutions of higher education as having very high research activity, located in the Southwestern United States. The researchers' affiliation as course instructors at this university at the time of data collection facilitated easier access to and selection of participants for the study. These participants were also selected due to a noted absence of instruments measuring engineering teaching efficacy beliefs for this demographic, and their crucial stage in needed professional development aligns with the NGSS's emphasis on integrating engineering in K-12 education. The participants ranged in age from 21 to 58 years (mean [M] = 28). Although they had varying exposure to college-level science courses (ranging from

1 to 10), none had taken an engineering course, highlighting the need for targeted educational measures in this area.

## Instrument

In the present work, STEBI-B, originally designed for pre-service teachers by Enochs and Riggs (1990), was adapted to develop the ETEBI. This adaptation was grounded in the proven adaptability and reliability of the STEBI across various teaching disciplines such as math (Enochs et al., 2000; MTEBI), environmental education (Sia, 1992; EEENI) and equitable science teaching for science teaching and learning for diverse learners (Ritter et al., 2001; SEBEST). Prior adaptations of the STEBI in different subject areas have successfully identified similar factor loadings, demonstrating its robustness and applicability in diverse educational contexts. In this sense, we used the original STEBI and modified it for the specific context of engineering education. Specifically, the STEBI instrument was adapted by replacing "science" with "engineering", "scientific inquiry" and "scientific experiments" with "engineering design process" or "engineering activities" and changing the overall context to focus on engineering teaching to create a self-efficacy tool specifically for pre-service teachers in engineering education. This modification provides an essential tool for assessing and enhancing the engineering teaching efficacy beliefs of future educators in this increasingly important science, technology, engineering, and mathematics (STEM) discipline. The ETEBI is composed of 23 items, each evaluated on a Likert scale ranging from strongly disagree (1) to strongly agree (5). This instrument is specifically designed to reflect engineering teaching efficacy beliefs. Sample items include "I will continually find better ways to teach engineering" and "I understand engineering concepts well enough to be effective in teaching elementary engineering". To provide evidence for the validity of the modified instrument, rigorous analyses including Rasch analysis and CFA were performed. The standards for educational & psychological testing (2014 edition) was used as a guideline for establishing the validity of the instrument (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014).

## Data Analysis

Prior to the analyses, an expert panel was consulted including six experts in the fields of science, engineering, STEM, and statistics to provide evidence for the face and content validity of the instrument. While an expert panel size exceeding ten members is considered optimal (Hyrkäs et al., 2003), a panel consisting of five to ten experts is also deemed acceptable (Kaya et al., 2023; Lynn, 1986), which aligns with the size of our panel. The experts were asked to rate each item in the instrument based on clarity, comprehensiveness, relevance, and appropriateness for engineering education context. The experts' agreement on the instrument items was assessed using Fleiss' kappa ($k$). The value was found to be [$k$=.72, $p<0.001$], which falls within the range of substantial agreement among the panelists (Landis & Koch, 1977; Xie et al., 2018). In other words, this value indicates strong consensus among the experts, suggesting that the instrument items are generally relevant and well-constructed for the engineering education context.

According to Bandura (1977), the self-efficacy construct comprises two dimensions: personal self-efficacy and outcome expectancy. Accordingly, these established theoretical factors were initially relied upon, while the researchers remained open to any revisions to the scale. In this study, using the Rasch model and confirmatory factor analyses, the construct validity and the factor structure of the model were analyzed. For the personal engineering teaching efficacy (PETE) scale, five items were worded positively (2, 5, 12, 18, 22), and eight items (3, 6, 8, 17, 19, 20, 21, 23) were worded negatively. For the Engineering teaching outcome expectancy (ETOE) subscale, eight items (1, 4, 7, 9, 11, 14, 15, 16) were worded positively and two items (10, 13) were worded negatively. The responses to each item on the scale were made on a Likert scale which ranges from 1 (strongly disagree) to 5 (strongly agree). The range of potential scores on the scale extends from a minimum of 23 to a maximum of 115. During the analyses, the items with negative phrasing underwent reverse coding so that all the scores were positively oriented for the data analyses.

### *Rasch analysis*

As part of the instrument validation process, the dataset was randomly divided, consisting of 561 pre-service elementary teachers, into two nearly equal groups to employ Rasch analysis and CFA. Specifically, 281 pre-service elementary teachers were selected for Rasch analysis, while the remaining 280 participants were used for CFA. Bond and Fox (2015) suggested that the Rasch model is a robust choice for the development and validation of Likert-type instruments, which are often used in pre-college STEM education research (Fisher, 1991). The study used the Rasch model over exploratory factor analysis (EFA) for analyzing the ETEBI Likert-type instrument data, as the Rasch model offers a more robust and nuanced approach, providing greater reliability and detailed insights into validity and uni-dimensionality compared to EFA. Rasch analysis employs the concept of item response theory (Embretson & Reise, 2013) that models the likelihood of an accurate response to an item on a scale as a function of the position of the item on the scale and the ability of the respondent (Boone & Rogan, 2005; Rasch, 1993). The ideal participant number for Rasch analysis is well established in the literature and around 250 is considered to be a sufficient number for the STEBI instrument (Linacre, 2024b). All the Rasch analyses were performed employing WINSTEPS (Linacre, 2024a) software version 5.2.5.1 in MS Windows 11 professional operating system. The study reported the infit and outfit indices, and person and item separations and reliability estimates suggested by Bond and Fox (2013). Infit is the measure of the model's ability to predict the responses of individuals, while outfit is the measure of the model's ability to predict the responses of a group of individuals. It is considered "acceptable" when the infit and outfit values fall between the ranges of 0.6 and 1.4 (Unfried et al., 2022; Wright, 1994; Wright & Linacre, 1994). Bond and Fox (2013) also suggested the optimum values for item separation, person separation, and item reliability. The person reliability estimates in the Rasch analysis are above 3.0, 2.0, 0.9, and 0.8 respectively. The Rasch analysis for both PETE and ETOE subscales of the ETEBI was performed in this study. The model was used to estimate the item and person parameters and to evaluate the alignment of the data with the model across the two subscales separately. The model was also utilized to identify any misfitting items, assess the separation reliability (Mallinson et al., 2004), and support the validity evidence of the instrument.

### *Confirmatory factor analysis*

CFA was then performed to analyze the construct validity of the resulting two-factor model, and further examine the resulting dimensional structure by using R programming software with the lavaan package. In CFA, it is critical to analyze the measures of close fit that are not dependent on the sample size. In that sense, the chi-square statistic, which is used to analyze the extent to which a proposed model differs from the data, is highly affected by sample size. In general, chi-square values are found to be higher for larger samples. Thus, to minimize the influence of sample size on chi-square results, the normed chi-square index ($\chi^2/df$) is suggested to be utilized to assess the model fit instead of standard $\chi^2$ values. Normed $\chi^2$ values equal to or below 5.0 are considered acceptable (Bentler, 1990). This study also reported the most commonly used CFA model fit measures, including the root mean squared error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). While the RMSEA measures the congruence between a hypothesized model and population data, CFI and TLI fit indices assess how well a hypothesized model fits as compared to the independence (null) model, which hypothesizes no correlation among the latent variables in the model. SRMR calculates the difference between observed and predicted correlations between model variables. An RMSEA value below 0.06 indicates a close to perfect fit (e.g., Hu & Bentler, 1999; Xia & Yang, 2019). The values of CFI and TLI greater than .90 and SRMR below .08 suggest a good and acceptable model (Schreiber et al., 2006).

## FINDINGS

### Rasch Analysis

While there were 10 items on the ETOE subscale initially, after performing the Rasch analysis iteratively, the study found that 2 items (items 10 and 13) did not fit the model (Wright & Masters, 1982). More particularly, infit and outfit values for those two items were above 1.50. Thus, these two items were excluded from the dataset and the Rasch analysis was subsequently reapplied in an iterative manner (refer to **Table 1**).

The final version of the instrument included 21 items in total (PETE subscale consists of 13 items and ETOE subscale consists of 8 items). The results of the Rasch

**Table 1.** Rasch analysis fit statistics–Infit and outfit mean square values

| Questionnaire items | SS | IMNSQ | OMNSQ |
|---|---|---|---|
| Q2. I will continually find better ways to teach engineering. | PETE | 1.18 | 1.13 |
| Q3. Even if I try very hard, I will not teach engineering as well as I will most subjects.* | PETE | 1.13 | 1.17 |
| Q5. I know the steps necessary to teach engineering design concepts effectively. | PETE | 1.11 | 1.10 |
| Q6. I will not be very effective in monitoring engineering design projects.* | PETE | .85 | .92 |
| Q8. I will generally teach engineering design ineffectively.* | PETE | .93 | .97 |
| Q12. I understand engineering design concepts well enough to be effective in teaching elementary engineering. | PETE | .97 | .96 |
| Q17. I will find it difficult to explain to students why engineering design projects are successful.* | PETE | .82 | .81 |
| Q18. I will typically be able to answer students' engineering questions. | PETE | .73 | .74 |
| Q19. I wonder if I will have the necessary skills to teach engineering.* | PETE | 1.15 | 1.16 |
| Q20. Given a choice, I will not invite the principal to evaluate my engineering design teaching.* | PETE | 1.32 | 1.27 |
| Q21. When a student has difficulty understanding an engineering design concept, I will usually be at a loss as to how to help the student understand it better.* | PETE | .70 | .69 |
| Q22. When teaching engineering, I will usually welcome student questions. | PETE | 1.20 | 1.09 |
| Q23. I do not know what to do to turn students on to engineering.* | PETE | .90 | .88 |

**Table 1 (Continued).** Rasch analysis fit statistics–Infit and outfit mean square values

| Questionnaire items | SS | IMNSQ | OMNSQ |
|---|---|---|---|
| Q1. When a student does better than usual in engineering activities, it is often because the teacher exerted a little extra effort. | ETOE | 1.31 | 1.31 |
| Q4. When the students do well in engineering design projects, it is often due to their teacher having found a more effective teaching approach. | ETOE | .82 | .78 |
| Q7. If students are underachieving in engineering design projects, it is most likely due to ineffective engineering teaching. | ETOE | .96 | 1.00 |
| Q9. The inadequacy of a student's engineering design background can be overcome by good teaching. | ETOE | 1.22 | 1.21 |
| Q11. When a low-achieving child makes progress in engineering, it is usually due to extra attention given by the teacher. | ETOE | 1.04 | 1.04 |
| Q14. The teacher is generally responsible for the achievement of students in engineering. | ETOE | .86 | .93 |
| Q15. Students' achievement in engineering is directly related to their teacher's effectiveness in engineering design teaching. | ETOE | .79 | .80 |
| Q16. If parents comment that their child is showing more interest in engineering at school, it is probably due to the performance of the child's teacher. | ETOE | .99 | 1.02 |

Note. SS: Subscale; IMNSQ: Infit MNSQ; OMNSQ: Outfit MNSQ; & *Reverse coded items

**Table 2.** Rasch analysis person and item separation and reliability estimates

| Subscales | Person separation | Reliability | Item separation | Reliability |
|---|---|---|---|---|
| PETE | 2.92 | .90 | 8.28 | .99 |
| ETOE | 2.07 | .81 | 5.45 | .97 |

analyses supported the uni-dimensionality of the two subscales of the ETEBI. More particularly, analysis of the two subscales of ETEBI with the Rasch method showed that both subscales had acceptable infit and outfit indices (Wright, 1994). The ETEBI has item separation estimates of more than 3.0 (strong) with reliability estimates of more than 0.9, and person separation estimates of more than 2.0 (good) with reliability estimates above 0.8 (Duncan et al., 2003). Therefore, the findings of this study suggest that both PETE and ETOE subscales are able to assess the constructs they are designed to measure (refer to **Table 2** and **Figure 1**). After the Rasch analysis, another set of data was used to run the CFA.

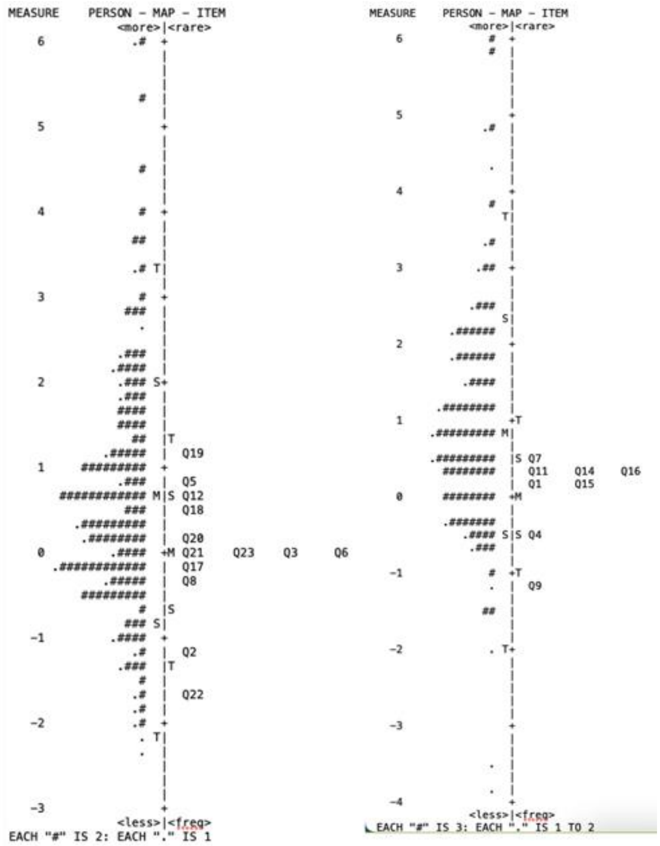


**Figure 1.** Wright maps for PETE and ETOE subscales (Source: Authors' own elaboration)

## Confirmatory Factor Analysis

Following the removal of two items from the ETOE subscale (Q10 and Q13) based on the Rasch analyses, CFA was conducted to investigate the extent to which the data conforms to the model and the factor structure of the final 21-item scale. The fit indices for the model (normed chi-square (Chi-square/df) = 1.64, *RMSEA* = 0.048, $p > .05$; *CFI* = 0.975; *TLI* = 0.972 and *SRMR* = 0.075) indicated that the two-factor model for the ETEBI is nearly a perfect fit, providing support for the construct validity of the ETEBI scale. **Figure 2** indicates the structure of the final ETEBI scale two-factor model based on the CFA.

A comprehensive psychometric test should also be accompanied by evidence of reliability. In that sense, Cronbach's coefficient alpha for each subscale was also measured to assess the internal consistency of the subscales. A Cronbach's alpha of 0.70 or higher indicates acceptable reliability (Cortina, 1993). Based on the analysis of Cronbach's alpha coefficients, the subscales were found to have high reliability (internal consistency): .897 for PETE, and .821 for ETOE. Therefore, the ETEBI is a valid and reliable instrument for assessing pre-service elementary teachers' beliefs about their efficacy in teaching engineering.

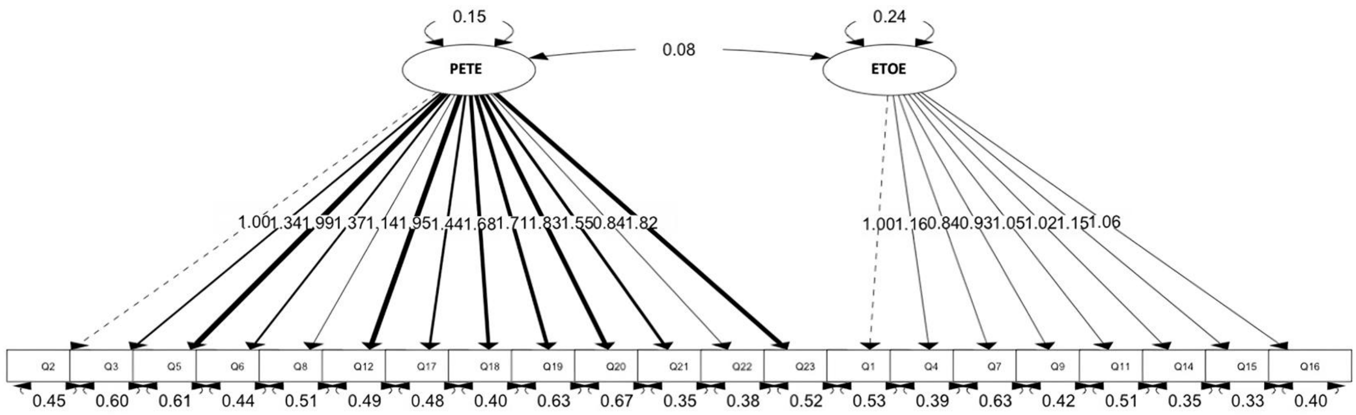


**Figure 2.** Path diagram of the CFA for a two-factor model with standardized loadings (Source: Authors' own elaboration)

## DISCUSSION AND IMPLICATIONS

The findings from this study, which focused on developing and providing evidence for the validity of the ETEBI, offer meaningful insights that align with and expand upon the existing literature on teacher self-efficacy and engineering education in elementary settings. The development of the ETEBI is a significant step in addressing the gap in measurement tools specifically designed for assessing the engineering teaching self-efficacy of pre-service elementary teachers. This aligns with the findings of Enochs and Riggs (1990) and subsequent studies, which emphasize the importance of teacher self-efficacy in influencing instructional practices and student outcomes. The ETEBI not only measures pre-service elementary teachers' engineering teaching efficacy beliefs but also provides a framework for identifying areas where additional support and professional development may be needed.

The current study highlights the complexity of teaching self-efficacy, particularly in engineering, a discipline not traditionally emphasized in elementary education. This resonates with the broader literature, which underscores the importance of subject-specific teaching self-efficacy (Bandura, 1997; Hammack & Ivey, 2017b). Targeting pre-service elementary teachers, the ETEBI responds to the growing focus on integrating engineering education in the K-12 curriculum as recommended by NGSS Lead States (2013).

Through the Rasch analysis, several items that did not fit the model were removed. Specifically, in the process of validating the ETEBI, items 10 and 13 were identified as misfitting and subsequently removed. This decision, guided by the principles of Rasch analysis, aligns with the broader literature on instrument development and the nuanced nature of self-efficacy constructs. Similar findings were noted in studies by Bleicher (2004) and Enoch and Riggs (1990), where certain items, especially items 10 and 13 in outcome expectancy scales were problematic due to cross-loading and low factor loadings. These studies highlight the importance of rigorous item analysis to ensure that each item accurately represents the intended construct. These misalignments could be due to the wording, negative wording, context-specific nuances, or the unique challenges and perceptions associated with teaching engineering at the elementary level (Enoch & Riggs, 1990; Unfried et al., 2022). It is important in educational measurement to ensure that items are not only theoretically sound but also contextually relevant. The removal of these items underscores the need for ongoing research and refinement of instruments like the ETEBI. Future studies might explore the specific reasons why these items did not fit well, possibly through qualitative methods or by testing alternative formulations of these items. Such investigations can provide deeper insights into the complex nature of self-efficacy beliefs in the domain of elementary engineering education.

The Rasch model and CFA as well as Cronbach's alpha provided robust evidence for the validity and reliability of the final ETEBI. This process of validation is crucial, as accurate measurement of self-efficacy beliefs is foundational for both educational research and practice. The findings of the CFA further support the construct validity of the ETEBI, aligning with the theoretical framework proposed by Bandura (1977, 1997) regarding the multi-dimensional nature of the self-efficacy construct.

As a strong variable in predicting teachers' classroom practices, teaching self-efficacy has been regarded as a noteworthy construct in teacher education. This study was undertaken to address the necessity of systematically assessing pre-service elementary teachers' engineering teaching self-efficacy. Building on the work of Enochs and Riggs (1990), this study provided ample evidence for the validity and reliability of the modified STEBI-B instrument in measuring pre-service elementary teachers' engineering teaching efficacy beliefs. The ETEBI, as an engineering content-specific instrument that is designed specifically for pre-service elementary teachers, enriches the existing pre-college engineering education literature and offers new insights into teacher self-efficacy in the context of engineering instruction in teacher education settings. The final PETE scale included 13 items and ETOE included 8 items excluding items 10 and 13. For further details on reverse coding and additional information, refer to Enochs and Riggs (1990).

It is crucial to lay a strong foundation for engineering education as early as elementary school because learners' ideas are forming at an early age. Elementary teachers should therefore be prepared to address this need. However, the necessity to assess the current status of pre-service teachers' engineering efficacy beliefs and/or the change in their engineering teaching efficacy beliefs following an intervention or a course has not been sufficiently met until now. The study contends that ample evidence was provided for future research to use the ETEBI instrument with confidence in assessing pre-service elementary teachers' PETE and ETOE beliefs. Diagnosing the status of pre-service elementary teachers' engineering teaching efficacy beliefs through such an instrument may allow us to make evidence-based inferences about the extent to which engineering instruction in their future classrooms will be aligned with student-centered contemporary practices advocated by the most recent national science and engineering teaching standards.

## CONCLUSION AND RECOMMENDATIONS

This study provided necessary evidence for the validity and reliability of the ETEBI by adapting the original STEBI-B (Enochs & Riggs, 1990) in measuring pre-service elementary teachers' engineering teaching efficacy beliefs. The study suggests that having this modified instrument readily available will further advance engineering education research, as future researchers can use it with peace of mind, confident in its psychometric qualities.

The study confidently claims that there is robust evidence for the reliability and validity of the ETEBI through the Rasch model and CFA as well as Cronbach's alpha. Therefore, individuals involved in research on engineering education, teacher training, and the development of educational policies can derive significant benefits from the data generated through the ETEBI because the ETEBI can enable teacher educators and education policymakers to efficiently evaluate the impact of any intervention or reform with regard to engineering education.

Moreover, the ETEBI might offer valuable data for future research, especially in investigating the relationship between teacher self-efficacy, instructional practices, and student outcomes in STEM education. Such research could be instrumental in shaping educational policies and practices that are responsive to contemporary educational needs.

The data were collected in one of the most diverse universities in the Southwestern United States. In this sense, there is a scholarly interest in evaluating the performance of the ETEBI in similar demographic settings or less diverse environments. Consequently, further research is strongly encouraged to investigate this aspect. Furthermore, future studies could benefit from incorporating more diverse samples and collecting detailed data on participants' backgrounds. In addition, employing qualitative methods to explore the reasons behind the misfitting items identified in the Rasch analysis would deepen the understanding of the variables influencing engineering teaching efficacy beliefs and could guide the refinement of the ETEBI for broader applicability.

### Limitations

Even though the study provided robust evidence for the validity and reliability of the ETEBI, several limitations must be acknowledged. First of all, the sample in this study predominantly included female pre-service elementary teachers (89.68%). However, it should be kept in mind that the population in this study is similar to the national statistics because most pre-service elementary teachers are female (DataUSA, 2024). Investigators who will utilize the ETEBI with higher percentages of male participants should keep this in mind. Second, even though the study was conducted in one of the most ethnically and racially diverse universities in the USA, data regarding the ethnic backgrounds of the participants were not collected. As mentioned earlier, it is of interest to understand how the ETEBI will perform across the racial and ethnic diversity in the country. Finally, the data were collected from a single institution, which may not necessarily be representative of pre-service teacher education institutions nationwide. Hence, future research should aim to include more diverse samples and multiple institutions to enhance the generalizability of the findings.

**Author contributions:** All authors collaboratively designed the research study, collected the data, and conducted the statistical analyses. Together, they analyzed the reliability and validity of the ETEBI. All authors have sufficiently contributed to the manuscript, and they have read and approved the final version.

**Funding:** No funding source is reported for this study.

**Ethical statement:** The authors stated that this study was reviewed and approved by the University of Nevada, Las Vegas (UNLV) Social/Behavioral Institutional Review Board (IRB) under Exempt Review on July 17, 2015, with the document number [780144-1]. The study was exempt under categories #1 and #2 according to Federal regulatory statutes 45 CFR 46.101(b). Informed consent was obtained from all participants, which included detailed information about the purpose of the study, procedures involved, potential risks and benefits, confidentiality of data, and the voluntary nature of participation. The collected data were kept private and secure, stored in password-protected and encrypted files, and were accessible only to the research team.

**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


AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Oxford University Press.

Albion, P. R., & Spence, K. G. (2013). Primary connections in a provincial Queensland school system: Relationships to science teaching self-efficacy and practices. *International Journal of Environmental and Science Education, 8*(3), 501-520.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Oxford University Press. https://doi.org/10.5860/choice.32-0900

Ashton, P. T., & Webb, R. B. (1986). Teachers' sense of efficacy, classroom behavior, and student achievement. In P. T. Ashton, & R. B. Webb (Eds.), *Teachers' sense of efficacy and student achievement* (pp. 125-144). Longman.

Ashton, P. T., Buhr, D., & Crocker, L. (1984). Teachers' sense of efficacy: A self- or norm-referenced construct? *Florida Journal of Educational Research, 26*(1), 29-41.

Ashton, P. T., Olejnik, S., Crocker, L., & McAuliffe, M. (1982, April). *Measurement problems in the study of teachers' sense of efficacy* [Paper presentation]. The 1982 Annual Meeting of the American Educational Research Association., New York.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037//0033-295x.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 national survey of science and mathematics education+*. Horizon Research, Inc. Retrieved July 20, 2024, from https://horizon-research.com/NSSME/wp-content/uploads/2020/04/Report_of_the_2018_NSSME.pdf

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research, Inc. Retrieved July 16, 2024, from https://eric.ed.gov/?id=ED541798

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), Article 238. https://doi.org/10.1037//0033-2909.107.2.238

Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring self-efficacy in preservice elementary teachers. *School Science and Mathematics, 104*(8), 383-391. https://doi.org/10.1111/j.1949-8594.2004.tb18004.x

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.

Boone, W., & Rogan, J. (2005). Rigour in quantitative analysis: The promise of Rasch analysis techniques. *African Journal of Research in Mathematics, Science and Technology Education, 9*(1), 25-38. https://doi.org/10.1080/10288457.2005.10740574

Boz, Y., & Cetin-Dindar, A. (2021). Teaching concerns, self-efficacy beliefs and constructivist learning environment of pre-service science teachers: A modelling study. *European Journal of Teacher Education, 46*(2), 274-292. https://doi.org/10.1080/02619768.2021.1919079

Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction, 66*, Article 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Calkins, L., Wiens, P., Parker, J., & Tschinkel, R. (2023). Teacher motivation and self-efficacy: How do specific motivations for entering teaching relate to teacher self-efficacy? *Journal of Education, 204*(2), 427-438. https://doi.org/10.1177/00220574221142300

Coladarci, T., & Fink, D. R. (1995, April). *Correlations among measures of teacher efficacy: Are they measuring the same thing?* [Paper presentation]. The annual meeting of the American Educational Research Association, San Francisco, CA.

Coladarci, T., & Fink, D. R. (1995). Correlations among measures of teacher efficacy: Are they measuring the same thing? In *Proceedings of the Annual Meeting of the American Educational Research Association*.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104. https://doi.org/10.1037//0021-9010.78.1.98

DataUSA. (2024). *Elementary & middle school teachers*. Retrieved July 16, 2024, from https://datausa.io/profile/soc/elementary-middle-school-teachers

Demir, S. (2020). The role of self-efficacy in job satisfaction, organizational commitment, motivation and job involvement. *Eurasian Journal of Educational Research, 20*(85), 205-224. https://doi.org/10.14689/ejer.2020.85.10

Duncan, P. W., Bode, R. K., Lai, S. M., Perera, S., & Glycine Antagonist in Neuroprotection Americas Investigators. (2003). Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Archives of Physical Medicine and Rehabilitation, 84*(7), 950-963. https://doi.org/10.1016/S0003-9993(03)00035-2

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press. https://doi.org/10.4324/9781410605269

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics, 90*(8), 694-706. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics, 100*(4), 194-202. https://doi.org/10.1111/j.1949-8594.2000.tb17256.x

Fisher, W. P. (1994). The Rasch debate: Validity and revolution in educational measurement. In M. Wilson (Ed.), *Objective measurement* (Vol. 2, pp. 36-72). Ablex.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*(4), 569-582. https://doi.org/10.1037//0022-0663.76.4.569

Guskey, T. R. (1981). Measurement of the responsibility teachers assume for academic successes and failures in the classroom. *Journal of Teacher Education, 32*(3), 44-51. https://doi.org/10.1177/002248718103200310

Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal, 31*(3), 627-643. https://doi.org/10.3102/00028312031003627

Hammack, R., & Ivey, T. (2017a). Elementary teachers' perceptions of engineering and engineering design. *Journal of Research in STEM Education, 3*(1/2), 48-68. https://doi.org/10.51355/jstem.2017.29

Hammack, R., & Ivey, T. (2017b). Examining elementary teachers' engineering self-efficacy and engineering teacher efficacy. *School Science and Mathematics, 117*(1-2), 52-62. https://doi.org/10.1111/ssm.12205

Henson, R. K. (2001). *Teacher self-efficacy: Substantive implications and measurement dilemmas* [Paper presentation]. The annual meeting of the Educational Research Exchange, College Station, TX.

Hsu, M. C., Purzer, S., & Cardella, M. E. (2011). Elementary teachers' views about teaching design, engineering, and technology. *Journal of Pre-College Engineering Education Research, 1*(2), Article 5.

Hu, L. T., & 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

Huang, S., Yin, H., & Lv, L. (2019). Job characteristics and teacher well-being: The mediation of teacher self-monitoring and teacher self-efficacy. *Educational Psychology, 39*(3), 313-331. https://doi.org/10.1080/01443410.2018.1543855

Hynes, M. M., Mathis, C., Purzer, S., Rynearson, A., & Siverling, E. (2017). Systematic review of research in P-12 engineering education from 2000-2015. *International Journal of Engineering Education, 33*(1), 453-462.

Hyrkäs, K., Appelqvist-Schmidlechner, K., & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of Nursing Studies, 40*(6), 619-625. https://doi.org/10.1016/s0020-7489(03)00036-1

Kang, E. J., Donovan, C., & McCarthy, M. J. (2018). Exploring elementary teachers' pedagogical content knowledge and confidence in implementing the NGSS science and engineering practices. *Journal of Science Teacher Education, 29*(1), 9-29. https://doi.org/10.1080/1046560x.2017.1415616

Karpudewan, M., Krishnan, P., Roth, W. M., & Ali, M. N. (2023). What research says about the relationships between Malaysian teachers' knowledge, perceived difficulties and self-efficacy, and practicing STEM teaching in schools. *The Asia-Pacific Education Researcher, 32*(3), 353-365. https://doi.org/10.1007/s40299-022-00658-1

Kaya, E., Deniz, H., & Yesilyurt, E. (2023). Toward developing a valid and reliable assessment of adults' nature of engineering views. *Journal of Engineering Education*, *112*(3), 634-673. https://doi.org/10.1002/jee.20524

Kaya, E., Newley, A., Deniz, H., Yesilyurt, E., & Newley, P. (2017). Introducing engineering design to a science teaching methods course through educational robotics and exploring changes in views of pre-service elementary teachers. *Journal of College Science Teaching*, *47*(2), 66-75. https://doi.org/10.2505/4/jcst17_047_02_66

Kaya, E., Newley, A., Yesilyurt, E., & Deniz, H. (2020). Measuring computational thinking teaching efficacy beliefs of preservice elementary teachers. *Journal of College Science Teaching*, *49*(6), 55-64. https://doi.org/10.1080/0047231x.2020.12290665

Kaya, E., Yesilyurt, E., Newley, A., & Deniz, H. (2019). Examining the impact of a computational thinking intervention on pre-service elementary science teachers' computational thinking teaching efficacy beliefs, interest and confidence. *Journal of Computers in Mathematics and Science Teaching, 38*(4), 385-392. Retrieved July 16, 2024 from https://www.learntechlib.org/primary/p/210970

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59-76. https://doi.org/10.1016/j.edurev.2014.06.001

Kuvac, M., & Koc, I. (2022). Enhancing preservice science teachers' perceptions of engineer and engineering through STEM education: A focus on drawings as evidence. *Research in Science & Technological Education, 41*(4), 1539-1559. https://doi.org/10.1080/02635143.2022.2052038

Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. *Biometrics, 33*(1), Article 159. https://doi.org/10.2307/2529310

Lazarides, R., Buchholz, J., & Rubach, C. (2018). Teacher enthusiasm and self-efficacy, student-perceived mastery goal orientation, and student motivation in mathematics classrooms. *Teaching and Teacher Education, 69*, 1-10. https://doi.org/10.1016/j.tate.2017.08.017

Linacre, J. M. (2024a). *A user's guide to Winsteps ministeps Rasch-model computer programs*. Retrieved July, 16, 2024, from https://www.winsteps.com/winman/copyright.htm

Linacre, J. M. (2024b). *Sample size and item calibration (or person measure) stability*. Institute for Objective Measurement. Retrieved July 16, 2024, from https://www.rasch.org/rmt/rmt74m.htm

Livers, S. D., Zelkowski, J., Harbour, K. E., McDaniel, S. C., & Gleason, J. (2019). An examination of the relationships of mathematics self-efficacy and teaching practices among elementary, secondary, and special education educators. *Investigations in Mathematics Learning, 12*(2), 96-109. https://doi.org/10.1080/19477503.2019.1670891

Lucero, M., Valcke, M., & Schellens, T. (2013). Teachers' beliefs and self-reported use of inquiry in science education in public primary schools. *International Journal of Science Education, 35*(8), 1407-1423. https://doi.org/10.1080/09500693.2012.704430

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-386. https://doi.org/10.1097/00006199-198611000-00017

Ma, K., Luo, J., Cavanagh, M., Dong, J., & Sun, M. (2023). Measuring teacher self-efficacy: Validating a new comprehensive scale among Chinese pre-service teachers. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.1063830

Mallinson, T., Stelmack, J., & Velozo, C. (2004). A comparison of the separation ratio and coefficient α in the creation of minimum item sets. *Medical Care, 42*(1), 1-17. https://doi.org/10.1097/01.mlr.0000103522.78233.c3

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology, 81*(2), 247-258. https://doi.org/10.1037//0022-0663.81.2.247

Mojavezi, A., & Tamiz, M. P. (2012). The impact of teacher self-efficacy on the students' motivation and achievement. *Theory and Practice in Language Studies, 2*(3), 483-491. https://doi.org/10.4304/tpls.2.3.483-491

NASEM. (2020). *Building capacity for teaching engineering in K-12 education*. National Academies Press. https://doi.org/10.17226/25612

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press.

Nie, Y., Tan, G. H., Liau, A. K., Lau, S., & Chua, B. L. (2012). The roles of teacher efficacy in instructional innovation: Its predictive relations to constructivist and didactic instruction. *Educational Research for Policy and Practice, 12*(1), 67-77. https://doi.org/10.1007/s10671-012-9128-y

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578. https://doi.org/10.3102/00346543066004543

Perera, H. N., & John, J. E. (2020). Teachers' self-efficacy beliefs for teaching math: Relations with teacher and student outcomes. *Contemporary Educational Psychology, 61*, Article 101842. https://doi.org/10.1016/j.cedpsych.2020.101842

Pleasants, J., & Olson, J. K. (2019). What is engineering? Elaborating the nature of engineering for K-12 education. *Science Education, 103*(1), 145-166. https://doi.org/10.1002/sce.21483

Purzer, Ş., Quintana-Cifuentes, J., & Menekse, M. (2022). The honeycomb of engineering framework: Philosophy of engineering guiding pre-college engineering education. *Journal of Engineering Education, 111*(1), 19-39. https://doi.org/10.1002/jee.20441

Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. MESA Press.

Ritter, J. M., Boone, W. J., & Rubba, P. A. (2001). Development of an instrument to assess prospective elementary teacher self-efficacy beliefs about equitable science teaching and learning (SEBEST). *Journal of Science Teacher Education, 12*(3), 175-198. https://doi.org/10.1023/a:1016747713585

Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcome. *The Journal of Educational Research, 74*(3), 185-190. https://doi.org/10.1080/00220671.1981.10885308

Ross, J. A., Cousins, J. B., Gadalla, T., & Hannay, L. (1999). Administrative assignment of teachers in restructuring secondary schools: The effect of out-of-field course responsibility on teacher efficacy. *Educational Administration Quarterly, 35*(5), 782-805. https://doi.org/10.1177/00131619921968824

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), Article 1. https://doi.org/10.1037/h0092976

Rubeck, M. E., & Enochs, E. G. (1991). *A path analytical model of variables that influence science and chemistry teaching self-efficacy and outcome expectancy in middle* *school science teachers* [Paper presentation]. The annual meeting of the National Association for Research in Science Teaching, Fontana, WI.

Savran, A., & Çakiroglu, J. (2003). Differences between elementary and secondary preservice science teachers' perceived efficacy beliefs and their classroom management beliefs. *Turkish Online Journal of Educational Technology, 2*(4), 15-20.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-338. https://doi.org/10.3200/joer.99.6.323-338

Sehgal, P., Nambudiri, R., & Mishra, S. K. (2017). Teacher effectiveness through self-efficacy, collaboration and principal leadership. *International Journal of Educational Management, 31*(4), 505-517. https://doi.org/10.1108/ijem-05-2016-0090

Shahzad, K., & Naureen, S. (2017). Impact of teacher self-efficacy on secondary school students' academic achievement. *Journal of Education and Educational Development, 4*(1), Article 48. https://doi.org/10.22555/joeed.v4i1.1050

Sia, A. P. (1992). *Preservice elementary teachers' perceived efficacy in teaching environmental education: A preliminary study* [Paper presentation]. The annual meeting of the North American Association for Environmental Education, Toronto.

Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *The Journal of Special Education, 27*(1), 66-81. https://doi.org/10.1177/002246699302700105

Temiz, T., & Topcu, M. S. (2013). Preservice teachers' teacher efficacy beliefs and constructivist-based teaching practice. *European Journal of Psychology of Education, 28*(4), 1435-1452. https://doi.org/10.1007/s10212-013-0174-5

Trygstad, P. J., Malzahn, K. A., Banilower, E. R., Plumley, C. L., & Bruce, A. D. (2020). Are all students getting equal access to high-quality science education? Data from the 2018 NSSME+. Horizon Research, Inc. Retrieved July 16, 2024, from https://horizon-research.com/NSSME/wp-content/uploads/2020/11/Equity_Science_FINAL.pdf

Trygstad, P. J., Smith, P. S., Banilower, E. R., & Nelson, M. M. (2013). *The status of elementary science education: Are we ready for the next generation science standards?* Horizon Research, Inc.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783-805. https://doi.org/10.1016/s0742-051x(01)00036-1

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248. https://doi.org/10.3102/00346543068002202

Unfried, A., Rachmatullah, A., Alexander, A. & Wiebe, E. (2022). An alternative to STEBI-A: Validation of the T-STEM science scale. *International Journal of STEM Education, 9*, Article 24. https://doi.org/10.1186/s40594-022-00339-x

van Uden, J. M., Ritzen, H., & Pieters, J. M. (2014). Engaging students: The role of teacher beliefs and interpersonal teacher behavior in fostering student engagement in vocational education. *Teaching and Teacher Education, 37*, 21-32. https://doi.org/10.1016/j.tate.2013.08.005

Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*, 370.

Wright, B. D., & Linacre, J. M. (1994). *The Rasch model as a foundation for the Lexile framework*. Retrieved July 16, 2024, from https://metametricsinc.com/research-publications/rasch-model-foundation-lexile-framework/

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*, 409-428. https://doi.org/10.3758/s13428-018-1055-2

Xie, Z., Gadepalli, C., & Cheetham, B. (2018). A study of chance-corrected agreement coefficients for the measurement of multi-rater consistency. *International Journal of Simulation: Systems, Science & Technology, 19*(2), 10.1-10.9. https://doi.org/10.5013/ijssst.a.19.02.10

Yesilyurt, E., Deniz, H., & Kaya, E. (2021). Exploring sources of engineering teaching self-efficacy for pre-service elementary teachers. *International Journal of STEM Education, 8*(1), Article 42. https://doi.org/10.1186/s40594-021-00299-8

Yoon Yoon, S., Evans, M. G., & Strobel, J. (2014). Validation of the teaching engineering self-efficacy scale for K-12 teachers: A structural equation modeling approach. *Journal of Engineering Education, 103*(3), 463-485. https://doi.org/10.1002/jee.20049

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being. *Review of Educational Research, 86*(4), 981-1015. https://doi.org/10.3102/0034654315626801