Validation of the Teaching Engineering Self-Efficacy Scale for K-12 Teachers: A Structural Equation Modeling Approach

So Yoon Yoon,[a] Miles G. Evans,[b] and Johannes Strobel[a]

[a]Texas A&M University, [b]Purdue University

Abstract

Background Teacher self-efficacy has received attention because of its direct relationship with teachers' classroom behaviors. Since engineering has been increasingly introduced in K-12 (precollege) education, development of an instrument to measure teachers' self-efficacy in the context of teaching engineering has been needed.

Purpose (Hypothesis) This study reports the development and validation of the Teaching Engineering Self-Efficacy Scale (TESS) for K-12 teachers.

Design/Method The items for the TESS were constructed through a comprehensive review of the literature regarding K-12 engineering education, the development of teachers' self-efficacy instruments in STEM areas, and K-12 teachers' reflections on integrating engineering into their classrooms. During the content and face validity process, we used structural equation modeling to identify and confirm the factor structure of the TESS, and used item-analyses for reliability evidence.

Results With data from 434 teachers in 19 states, exploratory and confirmatory factor analyses using structural equation modeling resulted in the TESS consisting of 23 items loading across four factors: engineering pedagogical content knowledge, engineering engagement, engineering disciplinary self-efficacy, and outcome expectancy. Cronbach's α ranged from 0.89 to 0.96 and exhibited high internal consistency reliability coefficients for the TESS.

Conclusions Teacher self-efficacy is a situation-specific construct because teachers' efficacy beliefs depend on the content area and teaching environment. Use of the TESS, as an instrument tailored for the engineering teaching context, can contribute to the literature on K-12 engineering education and improve the teaching of precollege engineering.

Keywords K-12 engineering education; teacher self-efficacy; TESS

Introduction

To teach engineering in K-12 (pre-college) classrooms means, for most teachers, to teach something for which they are not adequately prepared: pre-service teacher training does not require learning engineering, and there are no teaching licenses for engineering teaching (Katehi, Pearson, & Feder, 2009). There is, however, a large movement to provide in-service teachers with professional development to help them integrate engineering into their

  http://wileyonlinelibrary.com/journal/jee
July 2014, Vol. 103, No. 3, pp. 463–485 DOI 10.1002/jee.20049

classrooms (Capobianco, Diefes-Dux, & Mena, 2011; Cunningham, Lachapelle, & Keenan, 2010). Since; engineering is increasingly being introduced into K-12 education (Mathias-Riegel, 2001; Carr, Bennett IV, & Strobel, 2012), there is a strong need to develop an instrument to measure teachers' preparedness to undertake this subject.

A well-established construct for measuring teachers' preparedness and effect on student achievement is teacher self-efficacy towards teaching a specific discipline, which can be defined as the personal belief of a teacher in his or her ability to positively affect students' educational attainments in a particular discipline (Bandura, 1997). For example, teachers' self-efficacy in teaching mathematics significantly affects students' attainment in mathematics (Borko & Whitcomb, 2008). Instruments measuring teacher self-efficacy are context- and domain-specific (Bandura, 1997). In order to adequately address the needs of teachers and to evaluate the success of teacher professional development programs for K-12 engineering education, an instrument for measuring teacher self-efficacy in teaching engineering needs to be developed and rigorously tested.

Purpose of the Study

The purpose of this study was to rigorously develop, validate, and test the Teaching Engineering Self-efficacy Scale (TESS) as a tool for measuring teacher preparedness. For this study, we defined *teaching engineering self-efficacy* as a teacher's personal belief in his or her ability to positively affect students' learning of engineering. Because "there is no all-purpose measure of perceived self-efficacy" (Bandura, 2006, p. 307), we included various aspects of engineering in the context of teaching. Therefore, the TESS addresses aspects of engineering design, teamwork, connection to other subjects, and discipline issues in hands-on engineering activities. By exploring the responses on the TESS, researchers and educators will be able to investigate the dynamics between teachers and students in classroom engineering education.

Theoretical Framework

Self-efficacy is one's personal belief about one's capability to take an action toward an attainment (Bandura, 1977). The concept was introduced in Bandura's (1977) theory of social learning, and has been an important measurement tool in education ever since (Bandura, 1997). In particular, teacher self-efficacy has received attention from researchers because of findings that indicate its direct relationship with teachers' classroom behaviors, which in turn have a direct influence on students' performance (Coladarci, 1992; Gibson & Dembo, 1984; Muijs & Reynolds, 2002). For example, Gibson and Dembo (1984) revealed apparent differences in classroom behavior between high-efficacy and low-efficacy teachers. While low-efficacy teachers spent a lot of time in small-group instruction, high-efficacy teachers spent more time in whole-group instruction, monitoring and checking seatwork, and preparation. They also provided more praise per correct answer and less criticism per incorrect answer than did low-efficacy teachers. High-efficacy teachers also guided students to correct answers effectively through more questioning. Thus, high-efficacy teachers devoted more effort to teaching students, and did so with better instructional strategies than did low-efficacy teachers.

Researchers have consistently shown how students' psychological states were affected by teacher self-efficacy, a consequence of the relationship between teachers' self-efficacy and their commitment in class (Ashton & Webb, 1986; Midgley, Feldlaufer, & Eccles, 1989). In a two-year longitudinal study, Midgley, Feldlaufer, and Eccles (1989) showed how students' beliefs about their mathematical abilities changed according to the level of teaching

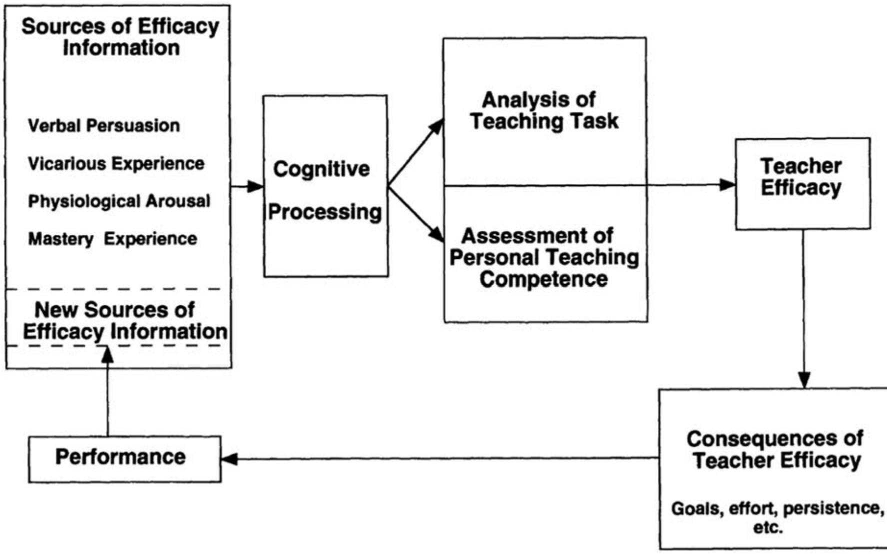

Figure 1 Framework of the teacher self-efficacy formation by Tschannen-Moran, Woolfolk Hoy, and Hoy (1998, p. 228).

mathematics self-efficacy of their teachers. Students who were taught by high-efficacy teachers in elementary schools reported a significant drop in expectancy of their own performance and achievement in middle school when they had low-efficacy middle school teachers; the drop in psychological states was more extreme for low-achieving than high-achieving students. The results indicate how much teachers' self-efficacy influences students' psychological states and their performance in class.

While reviewing studies on teachers' self-efficacy, Tschannen-Moran, Woolfolk Hoy, and Hoy (1998) conceptualized a framework to clarify confusions surrounding teacher self-efficacy. The framework outlines a cognitive procedure for the formation of teacher self-efficacy, which is situated in a feedback loop in nature (see Figure 1). Based on Bandura's (1986) four sources of self-efficacy (verbal persuasion, vicarious experience, physiological arousal, and mastery experience), this framework posits that the interaction between teachers' analysis of a teaching task and their self-assessment of teaching competence results in self-efficacy; self-efficacy shapes teachers' personal goals, amount of effort, and level of persistence in teaching students. Therefore, a teacher's performance in class is affected by his or her teaching self-efficacy, and, in turn, the outcome of his or her performance becomes the foundation of new sources of self-efficacy. Through this cycle, teacher self-efficacy is developed and changed. Here, note that teachers' appraisals of both tasks and their own teaching competence will differ by subject and environment. Thus, teacher self-efficacy varies by context and must be defined appropriately.

Existing literature provides evidence that perceived high self-efficacy by teachers plays a critical role in the functionality of classroom dynamics between teachers and students. Teachers' self-efficacy is not only related to their own behavior in class, but also to the goals, aspirations, and outcome expectancies they form for students; teacher self-efficacy thus has a great influence on students' self-efficacy, motivation, and achievement (Bandura, 1997;

Tschannen-Moran et al., 1998). In other words, teachers' self-efficacy functions as a moderator of their commitment and a mediator to predict student outcomes.

Several teacher self-efficacy instruments have been developed, validated, and utilized for various purposes in education (Tschannen-Moran & Woolfolk Hoy, 2001). At first, only general aspects of teacher self-efficacy were considered in the literature. Instruments were designed for all grade levels of teachers without discrimination between subject areas (Gibson & Dembo, 1984; Midgley et al., 1989). Reviewing more than 10 major teacher self-efficacy instruments in the literature, Tschannen-Moran et al. (1998) noted that most instruments were generally designed to assess global aspects of self-efficacy, so the instruments might not be useful tools for assessing specific aspects of self-efficacy. They concluded that an optimal level of specificity was necessary to make certain what is being measured by an instrument. In their later study, Tschannen-Moran and Hoy (2001) pointed out several problems with currently available instruments for measuring teacher self-efficacy. First, the validity and reliability of the instruments were still questionable. Second, the two-factor structure common in teacher self-efficacy instruments might be inadequate, because teacher self-efficacy is a complex phenomenon with many facets. Third, they pointed to a lack of consensus about to what extent instruments should be generalizable across different contexts as opposed to being context-specific.

Some context-specific models have been developed. Riggs and Enochs (1990) limited the content area of their instruments to the teaching of science to the elementary grade level when they developed their Science Teaching Efficacy Belief Instrument (STEBI). STEBI was designed to measure two constructs, outcome expectancy and self-efficacy, on the basis of Bandura's theoretical claim that behaviors are affected by both personal expectancy about the outcome and personal beliefs about teaching. Their use of a specific content area, in this case science, reflected the fact that teacher self-efficacy can vary depending on the content area. For example, while some teachers have high self-efficacy in teaching language arts, they may not have the same level in teaching science.

Since the first development of the STEBI, and due to its increasing use in science education, several variants have been developed and tested, each calibrated for a specific content area and different target population. The self-efficacy instruments for teaching mathematics for pre-service teachers (Enochs, Smith, & Huinker, 2000), chemistry for middle school teachers (Rubeck & Enochs, 1991), microcomputer utilization for in-service teachers (Enochs, Riggs, & Ellis, 1993), and STEM education for graduate teaching assistants (DeChenne & Enochs, 2010) are all rooted in the STEBI. Even though these instruments were developed on the basis of the STEBI, the types of constructs, the total number of items, and the phrasing of the statements in each item were tailored to fit the content and population targeted in each instrument. These modifications were necessary because an instrument that measures teacher self-efficacy needs enough sensitivity to address the self-efficacy situated in particular teaching contexts.

Teacher self-efficacy in the contexts of teaching science, mathematics, and technology has been investigated in various studies. But teachers' beliefs have rarely been explored in a K-12 engineering education setting since the introduction of engineering into precollege programs. K-12 teachers have rarely been exposed to engineering, are often unfamiliar with how to teach it, and must rely on pedagogical strategies familiar from other subjects (Rogers & Portsmore, 2004; Sun & Strobel, 2013). Thus, teacher training is a necessary prerequisite for the effective teaching of engineering. Under these circumstances, many professional development programs have been created to help teachers become qualified and confident in engineering and teaching engineering (Jeffers, Safferman, & Safferman, 2004).

Although recent efforts have been focused on teacher education, no valid and reliable instrument has been developed to measure teachers' self-efficacy in teaching K-12 engineering. Previous research on this matter was either conducted using generic teaching self-efficacy instruments (e.g., Holt, 2011) or with a qualitative research design (e.g., Hynes, 2012). Therefore, development of a psychometric instrument to assess teachers' beliefs in a quantifiable way is essential for further research in K-12 engineering education. Such an instrument can serve to clarify the belief systems of teachers who will teach engineering in class or who have already integrated engineering into their curricula, and will be beneficial for researchers and practitioners who are involved in teacher education programs; such an instrument will allow them to measure the effects of programs on increasing teachers' self-efficacy in teaching engineering.

Method

Instrument Development

Following the literature about scale development and psychometric testing procedures (Netemeyer, Bearden, & Sharma, 2003), we undertook several steps to develop an instrument to measure teachers' self-efficacy in teaching engineering, following guidelines set forth for self-efficacy instruments (Bandura, 2006; Tschannen-Moran et al., 1998). First, we reviewed studies that reported processes for developing and implementing new teacher self-efficacy instruments. The teacher self-efficacy instruments considered in this study included the Teacher Efficacy Scale (Gibson & Dembo, 1984), Science Teaching Efficacy Belief Instrument (Riggs & Enochs, 1990), Teacher Self-Efficacy Scale (Bandura, 2006), Ohio State Teacher Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001), and Teaching Technology Self-Efficacy (Teo, 2009). Table 1 gives additional information about these instruments. Studies using these instruments provided an idea of possible factor structures for the TESS.

Second, we reviewed the literature about the professional development programs for K-12 teachers' engineering education (Capobianco, 2011; Duncan, Diefes-Dux, & Gentry, 2011). This approach helped establish factors and refine items, so that the TESS would be situated in an engineering teaching context. Initial items and factors were modeled in detail in accordance with the reviewed instruments. Among the various factors that appeared in the teacher self-efficacy instruments in the literature, five were included:

Engineering content knowledge self-efficacy – teachers' personal belief in their knowledge of engineering that will be useful in a teaching context.

Instructional self-efficacy – teachers' personal belief in their ability to teach engineering to facilitate student learning.

Engagement self-efficacy – teachers' personal belief in their ability to engage students while teaching engineering.

Disciplinary self-efficacy – teachers' personal belief in their ability to cope with a wide range of student behaviors during engineering activities.

Outcome expectancy – teachers' personal belief in the effect of teaching on students' learning of engineering.

Third, we modified the existing items from the self-efficacy instruments in the literature and also generated new items to situate them in the context of teaching engineering. For consistency and clarification, item redundancies were eliminated and all items were rephrased

Table 1 Teaching Self-Efficacy Instruments in the Literature

Author(s)	Instrument	No. of Items	Scale	Constructs
Gibson & Dembo (1984)	Teacher Efficacy Scale	16	6-point Likert-type scale	Personal teaching efficacy General teaching efficacy
Riggs & Enochs (1990)	Science Teaching Efficacy Belief Instrument (STEBI)	25	5-point Likert-type scale	Personal science teaching efficacy (PSTE) Science teaching outcome expectancy (STOE)
Tschannen-Moran & Hoy (2001)	Ohio State Teacher Efficacy Scale	24	9-point Likert-type scale	Efficacy for instructional strategies Efficacy for classroom management Efficacy for student engagement
Bandura (2006)	Teacher Self-Efficacy Scale	30	9-point Likert-type scale	Instructional self-efficacy Disciplinary efficacy Influence on decision making Influence on school resources Enlisting parental involvement Enlisting community involvement Creating a positive school climate
Teo (2009)	Teaching Technology Self-Efficacy	16	7-point Likert-type scale	Basic teaching skills Advanced teaching skills Technology for pedagogy Traditional use of technology Constructivist use of technology

Note. Authors were ordered by the year of publication.

to be statements, not questions. While we observed negatively worded items in the instruments we reviewed, in the TESS all items are positively worded (e.g., "I can" instead of "I can't"). We decided against negatively worded items since prior research finds these may cause confusion for respondents (Netemeyer et al., 2003; Spector, 1992), produce cognitive burdens in processing information (Dillman, 2006), and distort reliability and validity evidence (Herche & Engellend, 1996). We also eliminated inconsistencies in word choice (e.g., using "student" instead of "child"). New items were also added to the initial item pool to fill in gaps, particularly for the engineering content knowledge self-efficacy construct, which is a new construct to measure teachers' pedagogical and content knowledge. In total, we generated 128 items grouped under the five factors for the next step of a content and face validity survey.

Fourth, all the items in the initial pool were judged by a panel of professors, current and previous K-12 teachers, and graduate students in the engineering and education disciplines. Six out of 14 panel members were previous or current teachers in K-12 education. Their feedback about the items was significant for checking appropriateness and clarity of the items and for improving them before collecting a large sample of teacher data. Fourteen panel members paired each item with a construct and indicated their level of confidence. Each item's score was the maximum number of people who agreed on a construct-item pairing with a high confidence level. If the scores of an item were relatively high for a specific construct, then that item was retained as a possible indicator of the construct. If the scores of an item spread across several constructs with relatively low scores, then it was discarded because it might not be a good indicator of one specific construct. Finally, reflecting the review and suggestions by the panel, 68 items were chosen to indicate the five factors.

Fifth, the format of the survey was determined using the suggestions for future improvements of teacher self-efficacy instruments (Boone, Townsend, & Staver, 2010; Tschannen-Moran & Woolfolk Hoy, 2001). For example, the level of scale was coded as a 6-point Likert-type scale (strongly disagree, moderately disagree, disagree slightly more than agree, agree slightly more than disagree, moderately agree, and strongly agree). We made this decision following Boone, Townsend, and Staver (2010), who conducted an experiment using the responses on the STEBI. Through reliability and Rasch analyses, they showed that the six-point response option, which does not have any neutral points or uncertainty points in the middle, provided better measurement properties than four- or five-point response scales.

Sample and Procedure

To evaluate both overall aspects and specific facets of teacher self-efficacy in teaching engineering, we selected K-12 teachers as the target population of the TESS. To reach teachers who taught or plan to teach engineering, we employed a snow-ball technique to maximize spread of the recruitment. We sent recruitment e-mail messages to (a) teachers we had taught before in national or regional workshops and conferences, (b) teacher professional development providers in the field of engineering education, whom we asked to forward the recruitment mail to their teacher participants, and (c) regional and national listservs serving teachers in K-12 education asking teachers who taught engineering to participate. Because we asked all the sources to forward the recruitment message to appropriate colleagues, we were able to determine neither the number of teachers our recruitment reached nor a proper response rate. In addition, we cannot provide information on where the teachers teach because we did not collect such information from the teachers.

Two Web-based survey programs, Qualtrics and SurveyMonkey, were used to construct the instrument online. Then teachers who were in K-12 education and who intended to incorporate or had already incorporated engineering into their classrooms were invited by e-mail to participate in the research and asked to respond to the TESS online. Teachers were also requested to fill out an online background survey to report their demographic information. In total, 434 participants completed the TESS, of which 12 failed to complete the demographic survey questions. Table 2 shows the distribution of participants by state. Table 3 gives the demographics of the participants. The majority of the teachers (89.6%) surveyed taught students in public schools (Table 4). The level of prior engineering exposure varied with 165 teachers (38.9%) reporting that they had attended a professional development program related to K-12 engineering education, and 269 teachers (70.0%) reporting that they had taught engineering in their classrooms. Teachers' ages ranged from 22 to 67, with $M = 43.8$ and $SD = 10.9$ ($n = 406$). A small minority of 28 teachers (6.5%) did not respond to the question asking about their age. On average, teachers ($n = 417$) spent 13.2 minutes to complete this initial version of the TESS, which included 68 items and the background survey. In terms of time, respondents who took more than one hour to complete the survey were considered as outliers.

Data Analysis

The distribution of responses on the 6-point Likert-type scale for each item was skewed and did not follow a normal distribution, so the maximum likelihood estimator, which assumes a normal distribution of responses, was not applicable for estimating parameters. Thus, the data were treated as categorical data, which are ordered and non-normal (Brown, 2006). To optimally estimate a factor structure of the underlying latent variables of the categorical data, factor analyses were completed using the Mplus 7.0 program (Muthén & Muthén, 2012).

The framework appropriate for conducting a factor analysis with categorical data differs from one appropriate for continuous data. In Mplus, robust weighted least-squares (WLSMV) is utilized as an estimator to obtain parameter estimates for the factor analysis. An exploratory factor analysis (EFA) was conducted with the data from the first round of data collection (involving 153 teachers) to investigate the underlying factor structures of the instrument and to identify irrelevant items that did not fit into any factors (Yoon, Evans, & Strobel, 2012). We carried out the EFA by calculating the polychoric correlation coefficients, eigenvalues, and factor loadings after oblique rotation of GEOMIN, which is the default rotation of the Mplus. After identifying the factor structure and items for the TESS, we conducted confirmatory factor analyses (CFAs) with the data from the second round of data collection (involving 281 teachers). Based on the fit indexes that Mplus 7.0 provides, the chi-square, root-mean-square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI) were used to judge CFA model fits (Brown, 2006; Grimm & Yarnold, 1995). We attempted several factor structure models with the items identified as the result of the EFA to refine the model fits of the CFAs using a structural equation modeling approach. As we finalized a factor structure and items for the TESS, we calculated the reliability coefficient of internal consistency, Cronbach's α, for each factor and for the single overall construct (teaching engineering self-efficacy) to investigate how items are interrelated within the factor.

Table 2 Teachers by Region and State

Region	n^a	Percentage
Northeast	41	9.4
CT	2	0.5
MA	2	0.5
NJ	4	0.9
NY	25	5.8
PA	8	1.8
Midwest	102	23.5
IL	1	0.2
IN	89	20.5
MI	1	0.2
OH	11	2.5
South	219	50.5
DE	1	0.2
MD	6	1.4
DC	4	0.9
FL	71	16.4
LA	7	1.6
TX	62	14.3
VA	68	15.7
West	53	14.5
CA	8	1.8
CO	30	6.9
HI	25	5.8
No response[b]	9	2.1
Total	434	100.0

Note. [a]Due to missing responses, the numbers are inconsistent with the total participant numbers. [b]While 12 teachers did not respond to the demographic survey, three teachers' state information could be identified. The data were sorted by region following Census Bureau-designated areas.

Results

Exploratory Factor Analysis Model

Factor extraction Since the data are ordered categorical variables, polychoric correlation coefficients among the 68 items were calculated. The correlation matrix indicated that the coefficients were all positively correlated; this correlation result meant that putative factors

Table 3 Demographic Information

Category	n^a	Percentage
Gender		
Female	333	76.7
Male	89	20.5
Non-respondent	12	2.8
Race/Ethnicity		
White	339	78.1
Hispanic	23	5.3
Black	19	4.4
Native Hawaiian or Other Pacific Islander	16	3.7
Asian	15	3.5
Multi-racial	7	1.6
American Indian or Alaska Native	3	0.7
Non-respondent	12	2.8
Age		
30 years or less	59	13.6
31–40 years	103	23.7
41–50 years	112	25.8
51–60 years	112	25.8
More than 60 years	20	4.6
Non-respondent	28	6.5
Full-time teaching experience		
5 years or less	84	19.4
6–10 years	88	20.0
11–20 years	160	36.9
21–30 years	65	15.0
31–40 years and over	25	5.5
Non-respondent	12	2.8
Teaching grade level		
Elementary school (K–G5)	273	62.9
Middle school (G6–G8)	80	18.4
High school (G9–G12)	68	15.7
Non-respondent	13	3.0
Engineering teacher professional development experience		
Yes	165	38.0
No	257	59.2
Non-respondent	12	2.8
Integration of engineering in class		
Yes	269	62.0
No	153	35.3
Non-respondent	12	2.8
Total	434	100.0

Note. [a]Due to incomplete responses, the numbers are inconsistent with the total participant numbers.

identified through an EFA are not independent. Since no correlations exceeded .85, multicollinearity was not observed, and hence no two items measure the same aspect of a construct, and each item contributes to a unique aspect of a factor.

The two criteria used to extract the number of factors underlying the data included the point of inflection of the curve in the scree plot (Cattell, 1966) and the number of eigenvalues greater than one (Kaiser, 1960). Following Kaiser's (1960) criteria, we retained factors with eigenvalues greater than one. This factor extraction yielded seven factors considered for inclusion in a putative factor structure for the TESS instrument.

Table 4 Teachers' School Demographics

Category[a]	n[b]	Percentage
School location		
Urban/City	153	35.3
Suburb	171	39.4
Town	45	10.4
Rural	53	12.2
School population		
Less than 500	118	27.2
500–1000	213	49.1
1000–2000	73	16.8
More than 2000	18	4.1
School type		
Public	389	89.6
Private	16	3.7
Magnet	8	1.8
Charter	3	0.7
Vocational	1	0.2
Total	434	100.0

Note. [a]The categories are guided by the classification of the National Center for Education Statistics. [b]Due to unspecified responses, the numbers are inconsistent with the total participant numbers.

Factor loadings Once a putative factor structure for the TESS was identified, the factor loadings of the items were gauged for each factor to decide which items constitute which factors. On the basis of Stevens' (2002) guideline about the relationship between the sample size and cutoff factor loading, items with a factor loading greater than 0.40 were considered significant for the designated factor. This cutoff functioned to suppress as irrelevant any items that did not fit well into the designated factor. If an item loaded onto more than one factor, then the item was excluded. No items were loaded onto the seventh factor with a value greater than the cutoff value of factor loading, so the seventh factor was not included in the final factor structure of the TESS. This analysis resulted in 41 items, out of the original 68, that fit into one of the six factors, as shown in Table 5. All 41 items had significant factor loadings onto one of six factors. In other words, each item had a unique contribution to one of the factors.

Construct match to factors Table 5 shows that the first 17 items, clustered on Factor 1, related to the construct *teacher self-efficacy in engineering content knowledge*. The three items on Factor 2 were associated with teachers' *motivational self-efficacy*, which is a new factor. Factor 3 consists of five items related to teachers' *instructional self-efficacy*. Factor 4 contains four items indicating teachers' *engagement self-efficacy*. Factor 5 includes six items constituting *disciplinary self-efficacy*. Finally, the last six items loaded on Factor 6 related to *outcome expectancy*.

Initial item and reliability analysis The overall reliability of the TESS with 41 items was Cronbach's $\alpha = 0.98$ from $n = 153$. Each construct housed in the TESS appeared to have good internal consistency as shown in Table 5. All items were worthy of retention because removal of any item would not increase Cronbach's α for any factor.

Confirmatory Factor Analysis Model

To confirm and refine the factor structure for the TESS, several CFAs were conducted with a new dataset from 281 teachers. We evaluated each CFA model following these steps:

Table 5 Exploratory Factor Analysis Results ($n = 153$)

	Item	Factor loading
Engineering content knowledge self-efficacy (Cronbach's α = 0.98)		
1	I can explain the different aspects of the engineering design process.	1.059[a]
2	I can discuss how given criteria affect the outcome of an engineering project.	1.028[a]
3	I can explain engineering concepts well enough to be effective in teaching engineering.	0.996
4	I can assess my students' engineering products.	0.974
5	I know how to teach engineering concepts effectively.	0.939
6	I can teach engineering as well as I do most subjects.	0.907
7	I can craft good questions about engineering for my students.	0.902
8	I can employ engineering activities in my classroom effectively.	0.839
9	I can discuss how engineering is connected to my daily life.	0.819
10	I can spend the time necessary to plan engineering lessons for my class.	0.808
11	I can explain the ways that engineering is used in the world.	0.775
12	I can describe the process of engineering design.	0.757
13	I can select appropriate materials for engineering activities.	0.721
14	I can create engineering activities at the appropriate level for my students.	0.702
15	I can stay current in my knowledge of engineering.	0.694
16	I can recognize and appreciate the engineering concepts in all subject areas.	0.650
17	I can guide my students' solution development with the engineering design process.	0.632
Motivational self-efficacy (Cronbach's α = 0.84)		
18	I can motivate students who show low interest in learning engineering.	0.755
19	I can increase students' interest in learning engineering.	0.661
20	Through engineering activities, I can make students enjoy the class more.	0.444
Instructional self-efficacy (Cronbach's α = 0.92)		
21	I can use a variety of assessment strategies for teaching engineering.	0.740
22	I can adequately assign my students to work at group activities like engineering.	0.702
23	I can plan engineering lessons based on each student's learning level.	0.681
24	I can gauge student comprehension of the engineering materials that I have taught.	0.679
25	I can help my students apply their engineering knowledge to real world situations.	0.550
Engagement self-efficacy (Cronbach's α = 0.88)		
26	I can promote a positive attitude toward engineering learning in my students.	0.690
27	I can encourage my students to think creatively during engineering activities and lessons.	0.596
28	I can encourage my students to think critically when practicing engineering.	0.517
29	I can encourage my students to interact with each other when participating in engineering activities.	0.498
Disciplinary self-efficacy (Cronbach's α = 0.94)		
30	I can control disruptive behavior in my classroom during engineering activities.	0.896
31	I can keep a few problem students from ruining an entire engineering lesson.	0.889
32	I can redirect defiant students during engineering lessons.	0.868
33	I can calm a student who is disruptive or noisy during engineering activities.	0.789
34	I can get through to students with behavior problems while teaching engineering.	0.569
35	I can establish a classroom management system for engineering activities.	0.542

Table 5 (continued)

	Item	Factor loading
Outcome expectancy (Cronbach's α = 0.88)		
36	I am generally responsible for my students' achievements in engineering.	0.638
37	When my students do better than usual in engineering, it is often because I exerted a little extra effort.	0.625
38	My effectiveness in engineering teaching can influence the achievement of students with low motivation.	0.505
39	When a student gets a better grade in engineering than he/she usually gets, it is often because I found better ways of teaching that student.	0.502
40	If I increase my effort in engineering teaching, I see significant change in students' engineering achievement.	0.471
41	I am responsible for my students' competence in engineering.	0.436

Note. [a]If categorical data are employed to indicate the latent factor structures, then factor loadings correspond to probit regression coefficients when WLSMV is employed. Thus, factor loadings greater than one are possible values.

Check the consistency of multiple goodness-of-fit indexes and judging the fit of the obtained CFA model to the data

Examine localized areas of poor fit, if any

Inspect parameter estimates, such as factor loadings, factor variances, and residual variances, to ensure the observed data's reliability on each item to the latent factor (Brown, 2006; Schreiber, Stage, King, Nora, & Barlow, 2006)

Table 6 shows some of the results from the attempted CFA models, including fit indexes, factors, and number of items. A CFA requires more constraints in the relationships between items and factors than the model identified though an EFA (Brown, 2006). Thus, as shown in Table 6, the TESS factor model was modified by removal of items that behaved poorly, and further changes resulted from the item elimination (Brown, 2006; Floyd & Widaman, 1995).

CFA model specification First, we applied the six-factor model (Model 1), which had resulted from the EFA, for the data on 41 items. All factor loadings were significant, and the CFI and TLI were in a good-fit range, defined as 0.95 and over; however, the RMSEA was not in an acceptable range, which is defined as 0.08 or less (Brown, 2006). In addition, we found high factor correlations over 0.85 among some factors, which indicates poor discriminant validity or overlapping constructs. Thus, a revision of the CFA model was necessary to optimize parsimony and improve interpretability.

CFA model revision The CFA model respecification was done in several ways on the basis of the factor correlations and the modification indexes (i.e., specific areas of the model misfit that show items with a discrepancy between the data and the proposed model). To acquire a parsimonious solution, we combined factors that showed the highest correlations, such as engineering content knowledge, motivational, and instructional self-efficacy, in sequence and eliminated items that cause large modification indexes.

Table 6 shows the changes that occurred in the model fit indexes, the number of factors, and the items in three successive CFA model respecifications: Model 2, Model 3, and Model 4. In Model 1, MS (motivational self-efficacy) consists of three items, but modification

Table 6 CFA Models with the Goodness-of-Fit Indexes ($n = 281$)

	Model 1	Model 2	Model 3	Model 4	Model 5 (Final model)
Fit Index					
Chi square	2557.84	1968.89	1036.86	600.26	600.57
df	764	655	395	244	226
p-value	< .001	< .001	< .001	< .001	< .001
RMSEA	0.091	0.084	0.076	0.077	0.077
90% CI	(0.088, 0.095)	(0.080, 0.089)	(0.070, 0.082)	(0.070, 0.085)	(0.069, 0.084)
CFI	0.967	0.975	0.982	0.986	0.986
TLI	0.976	0.093	0.980	0.984	0.984
No. of factors	6	5	5	4	5
No. of items	41	38	30	23	23
First-order factor	KS	KS	KS	KS	KS
	MS				
	IS	IS	IS		
	ES	ES	ES	ES	ES
	DS	DS	DS	DS	DS
	OE	OE	OE	OE	OE
Second-order factor					TES

Note. CI = confidence interval; KS = engineering pedagogical content knowledge self-efficacy; MS = engineering motivational self-efficacy; IS = engineering instructional self-efficacy; ES = engineering engagement self-efficacy; DS = engineering disciplinary self-efficacy; OE = engineering outcome expectancy; TES = teaching engineering self-efficacy; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

indices were large around those items, indicating areas of poor fit. Thus, MS items were not considered in Model 2; this exclusion resulted in better model fit indexes. The revision process from Model 2 to Model 4 took place because some items loading onto KS (engineering content knowledge self-efficacy) and IS (instructional self-efficacy) still caused large modification indexes, and the two constructs had a high factor correlation. Thus, problematic items were excluded in Model 3, and the items of IS were combined with those of KS (renamed as engineering pedagogical content knowledge self-efficacy hereafter). Finally, we arrived at Model 4, with four constructs indicated by 23 items. This model shows a significant degradation in RMSEA, to an acceptable range, as well as improved CFI and TLI compared with the initial model (Model 1).

Considering the fact that the four constructs are multifaceted aspects of the larger construct of teaching engineering self-efficacy (TES), and that we observed fairly high positive correlations among the four factors, we imposed a higher order factor on the model, using a structural equation modeling approach. Model 5 (the final model) shows that the second-order factor CFA model fits well with the data with all fit indexes similar to the ones in Model 4, $\chi^2(226) = 600.57$, $p < 0.01$, RMSEA = 0.077, CFI = 0.986, TLI = 0.984. Table 7 shows unstandardized and standardized parameter estimates of the final model, including factor loadings, variances, and residual variances, and Figure 2 reveals the second-order factor structure of the TESS created by the Mplus Diagrammer 1.0 (Muthén & Muthén, 2012).

Table 7 Parameter Estimates of the Final CFA Model (Model 5)

		Unstandardized		Standardized			
Factor (abbr.)	Item	Factor loading (R)	SE	Factor loading[b] (R)	SE	Factor variance (R^2)	Residual variance ($1\text{-}R^2$)
Engineering pedagogical content knowledge self-efficacy (KS)	2[a]	1.000	0.000	0.927	0.009	0.859	0.141
	4	1.013	0.009	0.940	0.008	0.884	0.116
	7	0.978	0.013	0.907	0.012	0.823	0.177
	8	1.015	0.011	0.941	0.008	0.885	0.115
	9	0.969	0.014	0.899	0.013	0.808	0.192
	10	0.765	0.028	0.710	0.028	0.504	0.496
	16	0.873	0.022	0.810	0.021	0.656	0.344
	17	1.011	0.009	0.938	0.008	0.880	0.120
	24	0.930	0.016	0.863	0.016	0.745	0.255
Engineering engagement self-efficacy (ES)	26[a]	1.000	0.000	0.877	0.015	0.769	0.231
	27	1.065	0.018	0.933	0.010	0.870	0.130
	28	1.034	0.017	0.907	0.012	0.823	0.177
	29	1.055	0.017	0.925	0.011	0.856	0.144
Engineering disciplinary self-efficacy (DS)	30[a]	1.000	0.000	0.916	0.012	0.839	0.161
	31	0.970	0.015	0.888	0.014	0.789	0.211
	33	0.857	0.023	0.784	0.022	0.615	0.385
	34	0.944	0.020	0.864	0.017	0.746	0.254
	35	1.016	0.014	0.931	0.011	0.867	0.133
Engineering outcome expectancy (OE)	36[a]	1.000	0.000	0.811	0.020	0.658	0.342
	37	0.926	0.025	0.751	0.025	0.564	0.436
	38	1.125	0.022	0.913	0.012	0.834	0.166
	39	0.909	0.028	0.737	0.026	0.543	0.457
	40	1.122	0.024	0.911	0.012	0.830	0.170
Teaching engineering self-efficacy (TES)	KS[a]	1.000	0.000	0.911	0.010	0.830	0.170
	ES	1.009	0.022	0.972	0.006	0.945	0.055
	DS	0.999	0.021	0.921	0.010	0.848	0.152
	OE	0.948	0.023	0.986	0.006	0.972	0.028

Note. [a]The item was used as a marker indicator to scale the latent factor, so the factor loading and standard error of the item were set to 1.0 and 0.0, respectively, as the default in M*plus* 7.0. [b]All 23 factor loadings are statistically significant with $p < 0.05$.

Final item and reliability analysis The four subscales of the TESS appeared to have good internal consistency, with a Cronbach's α ranging from 0.89 to 0.96: Cronbach's $\alpha = 0.96$ for engineering pedagogical content knowledge self-efficacy (KS), 0.93 for engineering engagement self-efficacy (ES), 0.92 for engineering disciplinary self-efficacy (DS), and 0.89 for engineering outcome expectancy (OE). The values of the corrected item-total correlation, which are correlations between each item and the total score on the TESS, ranged from 0.64 to 0.87, which indicates a good fit for items that correlate well with the total scale (Field, 2009). The overall reliability of the final 23-item TESS indicating teaching engineering self-efficacy (TES) was Cronbach's $\alpha = 0.98$ from $n = 281$. Again, all items were worthy of retention because removal of any item for each factor would not increase Cronbach's α.

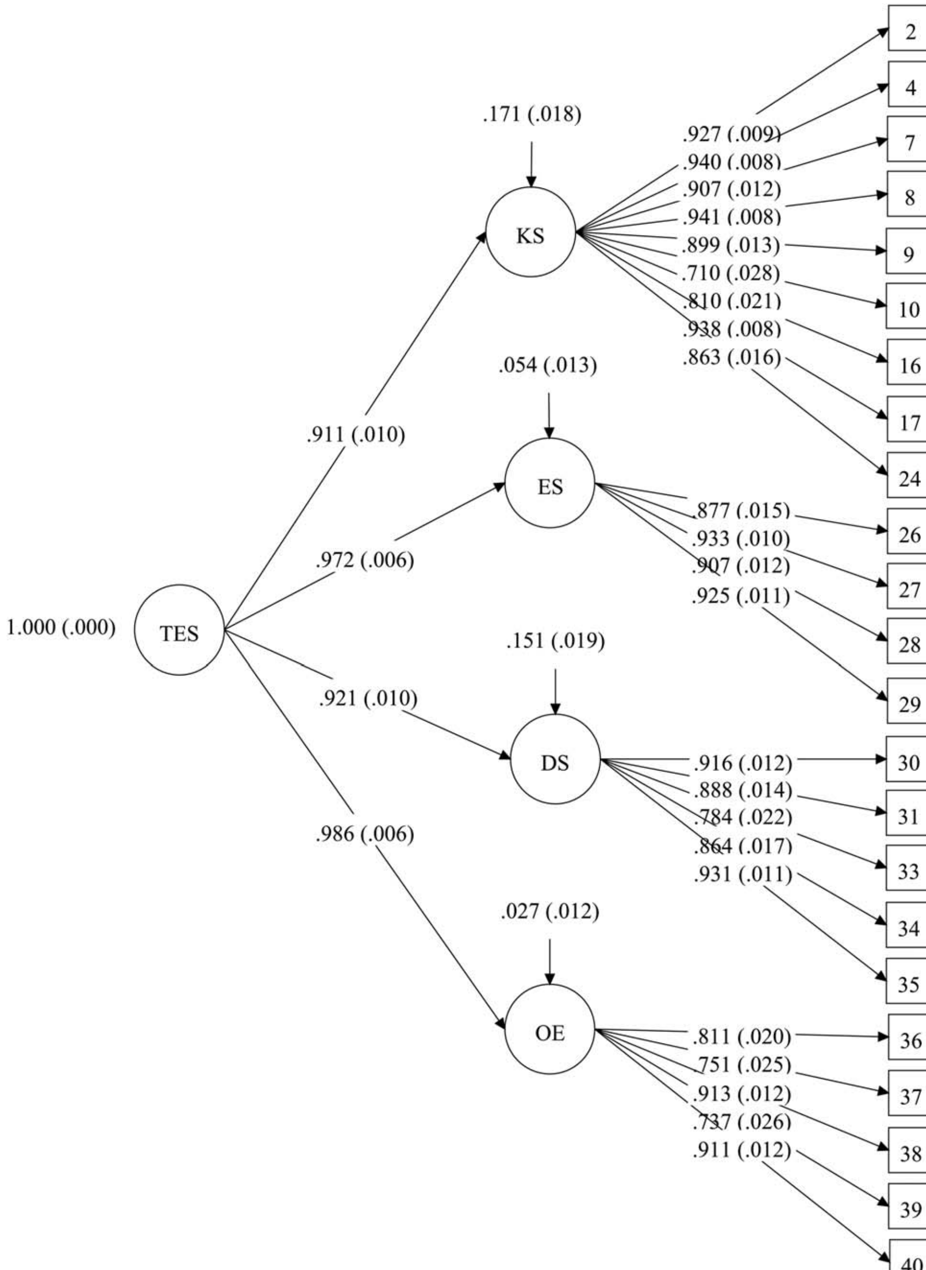

Figure 2 The final CFA model of the TESS with standardized estimates of factor loadings and standard errors in parentheses.

Discussion and Conclusion

The purpose of the study was to develop and validate the Teaching Engineering Self-Efficacy Scale (TESS) in order to provide an instrument to measure K-12 teachers' self-efficacy in teaching engineering. First, we conducted a literature review and identified five possible factors representing various aspects of self-efficacy: engineering content knowledge self-efficacy, instructional self-efficacy, engagement self-efficacy, disciplinary self-efficacy, and outcome expectancy. Then, we generated items to fit well with these constructs through a content and face validity process.

The EFA, using data from 153 teachers, resulted in six factors (engineering content knowledge self-efficacy, motivational self-efficacy, instructional self-efficacy, engagement self-efficacy, disciplinary self-efficacy, and outcome expectancy) significantly indicated by 41 items. Then the CFA, using data from 281 teachers, revealed that two factors identified from the EFA (engineering content knowledge self-efficacy and instructional self-efficacy) were highly correlated, so that the constructs seemed to have poor discriminant validity. Thus, for subsequent CFAs, we combined these two factors into a new construct that we named engineering pedagogical content knowledge self-efficacy. As a result of the EFA, we disaggregated the items for engagement self-efficacy into two constructs, which we renamed motivational self-efficacy and engagement self-efficacy. However, following the CFA, we excluded the items for the motivational self-efficacy because the large modification indexes indicated those items were the areas of poor fits and the overall model-fits without those items were improved. The items identified in the EFA as representing disciplinary self-efficacy and outcome expectancy were confirmed by the CFAs. The CFA data from this study yielded Cronbach's alphas ranging from 0.89 to 0.96 with an overall reliability coefficient of 0.98; this reliability evidence indicates that the TESS as used for K-12 teachers has good internal consistency.

In summary, through the EFA, we identified a factor structure of six factors represented by 41 items. Through the CFAs using structural equation modeling, we restructured the TESS and ultimately produced a model with four constructs represented by 23 items. The second-order CFA model supported the factor structure of the constructs of the TESS. This result implies that the four subscales, engineering pedagogical content knowledge self-efficacy (KS), engineering engagement self-efficacy (ES), engineering disciplinary self-efficacy (DS), and engineering outcome expectancy (OE), contribute to a single overall construct, the teaching engineering self-efficacy (TES). Summing scores of all items into a single score is a reasonable way to indicate the degree of TES. Table 8 contains the definition for each construct and the overall definition for the teaching engineering self-efficacy (TES) construct. In Appendix A, we present the final version of the TESS, with 23 items listed in order of the constructs to aid a logical flow of thought as respondents go through the survey. Directions for scoring the TESS follow in Appendix B.

Limitations and Future Studies

Science and mathematics have a long history in the K-12 schooling system and are well integrated into the preparation and continuous training of teachers. Consequently, many science and math concepts are shared among educators. That engineering in K-12 is yet to be fully conceptualized (Katehi, Pearson, & Feder, 2009; National Academy of Engineering, 2010), not only affects the practice of teaching engineering in K-12, but has additional consequences for this study. A limitation of this study is that it presumes the participating members of

Table 8 Constructs of the Teaching Engineering Self-Efficacy Scale (TESS)

Construct	Abbreviation	Definition
Engineering pedagogical content knowledge self-efficacy	KS	Teachers' personal belief in their ability to teach engineering to facilitate student learning, based on knowledge of engineering that will be useful in a teaching context.
Engineering engagement self-efficacy	ES	Teachers' personal belief in their ability to engage students while teaching engineering.
Engineering disciplinary self-efficacy	DS	Teachers' personal belief in their ability to cope with a wide range of student behaviors during engineering activities.
Engineering outcome expectancy	OE	Teachers' personal belief in the effect of teaching on student learning of engineering.
Teaching engineering self-efficacy	TES	Teachers' personal belief in their ability to positively affect students' learning of engineering that reflects the multifaceted nature of self-efficacy of teaching engineering.

the K-12 teaching community share a definition of engineering that they had in mind while answering the survey. However, the same definition might not be shared by all members of the community or by all participants of the study. Our comprehensive literature review addresses some of the concerns, yet future research on the effect of different conceptualizations of engineering in K-12 and their influence on teachers' self-efficacy is necessary.

The target users of the TESS are the K-12 teachers throughout the United States. The sampling for this study, though, is not random, and the data consist of teachers in only 19 states. In future research studies, we will continue to try to broaden the sample of the population that we use. As well, the consistency of the validity across different teaching grades has not yet been examined. The results of this study, then, are limited in terms of their generalizability beyond the sample characteristics of this study. Thus, for future research, conducting CFAs with different datasets will assure finalization of the items and factor structure of the TESS.

Item analyses based on classical test theory and item response theory with a large dataset will reveal overall psychometric properties of the TESS, such as individual item and overall test characteristics. A multiple-groups CFA will check measurement invariance of the TESS by examining the factor structure of the items of the instrument across different subgroups. This process will test for any bias in the TESS against a specific gender or teaching grade level. Additional evaluations of validity, such as convergent, discriminant, concurrent, and predictive, will facilitate researchers' understanding of the constructs measured by the TESS, such as how each construct is correlated with different measures (Anastasi & Urbina, 1997; Gregory, 1996). In sum, this effort will lead to the accumulation of knowledge on the psychometric properties of the TESS. The goal is creation of a TESS technical manual to guide best practices in use of the instrument by researchers and educational practitioners.

Significance of the Study

Teacher self-efficacy is a situation-specific construct because teacher efficacy beliefs depend on the content area and teaching environment (Bandura, 1997). Thus, the use of the TESS, as a teacher self-efficacy instrument tailored for the engineering teaching context, is expected to contribute to the literature on K-12 engineering education. First, the TESS can easily serve to diagnose and clarify the teacher's self-efficacy system and lead to further understanding of teachers' behavior in class. Second, when preparation of teachers occurs through in-service, pre-service, or professional development programs, the instrument allows researchers to examine teachers' beliefs, attitudes, and behavior patterns upon entering such programs, and then to assess how the programs have changed them. Thus, the TESS can be used as one evaluation tool for teacher preparation programs. Third, after diagnosing the current status of teachers' self-efficacy, the measure will be beneficial in helping trainers determine the best approaches to increase the self-efficacy of teachers according to which construct area they are weakest in. For example, teachers with low efficacy in instructional strategies may need different approaches in training from teachers with low efficacy in classroom disciplinary problems. Fourth, researchers using the TESS can extend the investigation of the relationship between teachers' self-efficacy and students' achievement, as situated in teaching and learning engineering, while also considering other plausible factors that may affect teachers' behavior in class and students' performance. In conclusion, we expect that the TESS can lead to diverse approaches in research on and training in K-12 engineering education.

Acknowledgments

This article was based in part on a paper presented at the 2012 American Society of Engineering Education Annual Conference, San Antonio, Texas (Yoon, Evans, & Strobel, 2012).

References

Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7th ed). Upper Saddle River, NJ: Prentice-Hall.

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York, NY: Longman.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

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

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

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Adolescence and education: Vol. 5. Self-efficacy beliefs of adolescents* (pp. 307–337). Charlotte, NC: IAP-Information Age Publishing.

Boone, W. J., Townsend, J. S., & Staver, J. (2010). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education*, *95*, 258–280. doi:10.1002/sce.20413

Borko, H., & Whitcomb, J. A. (2008). Teachers, teaching, and teacher education: Comments of the National Mathematics Advisory Panel's Report. *Educational Researcher*, *37*, 565– 572. doi:10.3102/0013189X08328877

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

Capobianco, B. M. (2011). Exploring a science teacher's uncertainty with integrating engineering design: An action research study. *Journal of Science Teacher Education*, *22*, 645–660. doi: 10.1007/s10972-010-9203-2

Capobianco, B. M., Diefes-Dux, H. A., & Mena, I. B. (2011). Elementary school teachers' attempts at integrating engineering design: Transformation or assimilation? *Proceedings of* *the ASEE Annual Conference and Exposition*, Vancouver, Canada. http://www.asee.org/public/conferences/1/papers/495/view

Carr, R. L., Bennett IV, L. D., & Strobel, J. (2012). Engineering in the K-12 STEM standards of the 50 U.S. states: An analysis of presence and extent. *Journal of Engineering Education*, *101*, 539–564. doi:10.1002/j.2168–9830.2012.tb00061.x

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245–276. doi:10.1207/s15327906mbr0102_10

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, *60*, 323–337. doi:10.1080/00220973.1992.9943869

Cunningham, C. M., Lachapelle, C. P., & Keenan, K. (2010). *Elementary teachers' changing ideas about STEM and STEM pedagogy through interaction with a pedagogically supportive STEM curriculum*. Paper presented at the K-12 Engineering and Design Education Research Summit, Seaside, OR.

DeChenne, S. E., & Enochs, L. (2010). *Measuring the teaching self-efficacy of science, technology, engineering, and math graduate teaching assistants*. Paper presented at the 2010 American Educational Research Association Annual Conference, Denver, CO.

Dillman, D. A. (2006). *Mail and internet surveys: The tailored design method* (2nd ed.). New York, NY: John Wiley & Sons.

Duncan, D., Diefes-Dux, H. A., & Gentry, M. (2011). Professional development through engineering academies: An examination of elementary teacher recognition and understanding of engineering. *Journal of Engineering Education*, *100*, 520–539. doi:10.1002/j.2168–9830.2011.tb00025.x

Enochs, L. G., Riggs, I. M., & Ellis, J. D. (1993). The development and partial validation of microcomputer utilization in teaching efficacy beliefs instrument in a science setting. *School Science and Mathematics*, *93*, 257–263.

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics*, *100*, 194–202. doi:10.1111/j.1949–8594.2000.tb17256.x

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*, 286–299. doi:10.1037/1040–3590.7.3.286

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

Gregory, R. J. (1996). *Psychological testing: History, principles, and applications* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Grimm, L. G., & Yarnold, P. R. (Eds.) (1995). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.

Herche, J., & Engellend, B. (1996). Reversed polarity items and scale dimensionality. *Journal of the Academy of Marketing Science*, *24*, 366–374. doi:10.1177/0092070396244007

Holt, B. D. (2011). *An exploratory study of Project Lead the Way secondary engineering educators' self-efficacy* (Doctoral dissertation). Virginia Polytechnic Institute and State University.

Hynes, M. (2012). Middle-school teachers' understanding and teaching of the engineering design process: A look at subject matter and pedagogical content knowledge. *International Journal of Technology and Design Education*, *22*, 345–360. doi:10.1007/s10798-010-9142-4

Jeffers, A. T., Safferman, A. G., & Safferman, S. I. (2004). Understanding K–12 engineering outreach programs. *Journal of Professional Issues in Engineering Education and Practice*, *130*, 95–108. doi:10.1061/(ASCE)1052–3928(2004)130:2(95)

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. doi:10.1177/001316446002000116

Katehi, L., Pearson, G., & Feder, M. (Eds.). (2009). *Engineering in K-12 education: Understanding the status and improving the prospects*. Committee on K-12 Engineering Education, National Academy of Engineering and National Research Council. Washington, DC: National Academies Press.

Mathias-Riegel, B. (2001). Engineering that's elementary. *Prism*, *10*(7), 34–36.

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*, 247–258. doi:10.1037/0022–0663.81.2.247

Muijs, D., & Reynolds, D. (2002). Teachers' beliefs and behaviors: What really matters. *Journal of Classroom Interaction*, *37*, 3–15.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus: Statistical analysis with latent variables: User's guide* (7th ed.). Los Angeles, CA: Authors.

National Academy of Engineering. (2010). *Standards for K-12 engineering education?* Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=12990.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures*. Thousand Oaks, CA: Sage.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*, 625–637. doi:10.1002/sce.3730740605

Rogers, C., & Portsmore, M. (2004). Bringing engineering to elementary school. *Journal of STEM Education: Innovations and Research*, *5*, 17–28.

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 presented at the annual meeting of the National Association for Research in Science Teaching, Fontana, WI.

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*, *99*, 323–337. doi:10.3200/JOER.99.6.323–338

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Lawrence Erlbaum.

Sun, Y., & Strobel, J. (2013). Elementary Engineering Education (EEE) adoption and expertise development framework: An inductive and deductive study. *Journal of Pre-College Engineering Education Research*, *3*(1), Article 4. Retrieved from http://docs.lib.purdue.edu/jpeer/vol3/iss1/4/

Teo, T. (2009). Examining the relationship between student teachers' self-efficacy beliefs and their intended uses of technology for teaching: A structural equation modeling approach. *Turkish Online Journal of Educational Technology*, *8*(4), 7–16. Retrieved from http://www.tojet.net/articles/v8i4/841.pdf

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*, 783–805.

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

Yoon, S. Y., Evans, M. G., & Strobel, J. (2012). Development of the Teaching Engineering Self-Efficacy Scale (TESS) for K–12 teachers. *Proceedings of the ASEE Annual Conference and Exposition*, San Antonio, TX.

Authors

*So Yoon Yoon is currently a post–doctoral research associate at Engineering Academic and Student Affairs, Dwight Look College of Engineering, Texas A&M University, College Station, TX, 77843; soyoon@tamu.edu.

Miles G. Evans is a junior in the environmental and ecological engineering program, College of Engineering, Purdue University, Potter Engineering Center, 500 Central Drive, West Lafayette, IN 47907–2022; evans93@purdue.edu.

Johannes Strobel is Director of the Educational Outreach Programs and an associate professor of engineering technology and industrial distribution and teaching, learning & culture in the College of Education, Texas A&M University, Zachry Hall of Engineering, College Station, TX, 77843; jstrobel@tamu.edu.

*corresponding author

Appendix A
Teaching Engineering Self-Efficacy Scale (TESS)

Directions: This survey contains statements about teachers' teaching engineering self-efficacy. Here, teaching engineering self-efficacy is defined as teachers' personal belief in their teaching engineering ability to positively affect student learning of engineering. Please indicate the degree to which you agree or disagree with each statement below by marking on the appropriate number to the right of each statement.

1 = Strongly Disagree
2 = Moderately Disagree
3 = Disagree slightly more than agree
4 = Agree slightly more than disagree
5 = Moderately agree
6 = Strongly agree

1.	I can discuss how engineering is connected to my daily life.	1	2	3	4	5	6
2.	I can recognize and appreciate the engineering concepts in all subject areas.	1	2	3	4	5	6
3.	I can spend the time necessary to plan engineering lessons for my class.	1	2	3	4	5	6
4.	I can employ engineering activities in my classroom effectively.	1	2	3	4	5	6
5.	I can craft good questions about engineering for my students.	1	2	3	4	5	6
6.	I can discuss how given criteria affect the outcome of an engineering project.	1	2	3	4	5	6
7.	I can guide my students' solution development with the engineering design process.	1	2	3	4	5	6
8.	I can gauge student comprehension of the engineering materials that I have taught.	1	2	3	4	5	6
9.	I can assess my students' engineering products.	1	2	3	4	5	6
10.	I can promote a positive attitude toward engineering learning in my students.	1	2	3	4	5	6
11.	I can encourage my students to think critically when practicing engineering.	1	2	3	4	5	6
12.	I can encourage my students to interact with each other when participating in engineering activities.	1	2	3	4	5	6
13.	I can encourage my students to think creativity during engineering activities and lessons.	1	2	3	4	5	6
14.	I can calm a student who is disruptive or noisy during engineering activities.	1	2	3	4	5	6
15.	I can get through to students with behavior problems while teaching engineering.	1	2	3	4	5	6
16.	I can keep a few problem students from ruining an entire engineering lesson.	1	2	3	4	5	6
17.	I can control disruptive behavior in my classroom during engineering activities.	1	2	3	4	5	6
18.	I can establish a classroom management system for engineering activities.	1	2	3	4	5	6
19.	When a student gets a better grade in engineering than he/she usually gets, it is often because I found better ways of teaching that student.	1	2	3	4	5	6
20.	When my students do better than usual in engineering, it is often because I exerted a little extra effort.	1	2	3	4	5	6
21.	If I increase my effort in engineering teaching, I see significant change in students' engineering achievement.	1	2	3	4	5	6
22.	I am generally responsible for my students' achievements in engineering.	1	2	3	4	5	6
23.	My effectiveness in engineering teaching can influence the achievement of students with low motivation.	1	2	3	4	5	6

Appendix B
Directions for Scoring the Teaching Engineering Self-Efficacy Scale (TESS)

There are two ways of scoring the TESS: scoring a raw mean score of each construct and the overall raw score of self-efficacy in teaching engineering.

Method 1 assesses a teacher's self-efficacy in one of the four constructs that the TESS is designed to measure. This is done by first computing the unweighted means of a teacher's score on the items that load on each construct (subscale factor). Table B1 matches each item to a construct.

Table B1 Constructs and Corresponding Items that Constitute Teaching Engineering Self-Efficacy (TES)

Construct (Subscale factor)	Abbreviation	No. of items	Items
Engineering pedagogical content knowledge self-efficacy	KS	9	1–9
Engineering engagement self-efficacy	ES	4	10–13
Engineering disciplinary self-efficacy	DS	5	14–18
Engineering outcome expectancy	OE	5	19–23

For example,

The mean KS score = (sum of scores from items, 1−9) / 9.

The mean ES score = (sum of scores from items, 10−13) / 4.

The mean DS score = (sum of scores from items, 14−18) / 5.

The mean OE score = (sum of scores from items, 19−23) / 5.

The TESS is a measure on a 6-point Likert-type scale (1 = *strongly disagree*, 2 = *moderately disagree*, 3 = *disagree slightly more than agree*, 4 = *agree slightly more than disagree*, 5 = *moderately agree*, and 6 = *strongly agree*), so the score for each item ranges from 1 to 6. Thus, the mean score for each construct has the same range.

Method 2 computes the overall self-efficacy in teaching engineering (TES) of a teacher. This is done by first completing Method 1 for each construct and then summing each of the scores. Thus, the maximum TES score is 24 and the minimum is 4.

For example,

The overall TES score = sum of all the four mean construct scores.