Journal of Research in STEM Education
Vol 10, No 1-2, December 2024, 3-18
https://doi.org/10.51355/j-stem.2024.136

RESEARCH REPORT

The Development and Validation of the Children's Engineering Self-Efficacy Scale (CESES)

Amy Catalano[1]
Hofstra University, USA

Abstract
In order to best prepare future generations to solve societal challenges, students should be provided with a foundation in problem solving skills. Accordingly, the Next Generation Science Standards (NGSS) incorporates engineering design process skills along with the scientific practices. While engineering curriculum opportunities have expanded, research testing the efficacy of engineering design instruction on problem solving skills is limited. Self-efficacy research has long drawn connections of high science self-efficacy to improved attitudes and performance. However, little research has examined engineering design self-efficacy, and none has been conducted with children. This study reports the development and validation of the Children's Engineering Self-Efficacy Scale (CESES). Participants included 212 children in grades 3-7 from a variety of instructional backgrounds. Initial results showed that the instrument produced a Cronbach's alpha of .81. Factor analysis resulted in a five-factor model explaining 56.72 % of the variance. Additional analyses showed that the majority of participants had good to high engineering self-efficacy. The result of this research has implications for examining the growth of self-efficacy after STEM or engineering interventions both in school programs and in out-of-school programs.

Keywords: *Engineering design process, Self-efficacy, STEM education, surveys and questionnaires, scale validation*

The ability to problem solves is an integral life and academic skill. In the last decade educational reform has emphasized the development of 21st century skills, which are the skills modern students require to function effectively in a technologically based society (Larson & Miller, 2011; National Science Teacher Association [NSTA], 2011). Twenty-first century skills include media and digital literacy, creative thinking and problem solving, among other skills. Accordingly, the Next Generation Science Standards (NGSS) emphasize these skills with the inclusion of the engineering design practices (NSTA, 2011). "Providing students a foundation in

---

[1] Hagedorn Hall 0251, Hempstead, N.Y. 11549-1000, USA, amy.catalano@hofstra.edu

To cite this article: Catalano, A. (2024). The development and validation of the children's engineering self-efficacy scale (CESES). *Journal of Research in STEM Education, 10*(1-2), 3-18. https://doi.org/10.51355/j-stem.2024.136

ISSN: 2149-8504 (online)

engineering design allows them to better engage in and aspire to solve the major societal and environmental challenges they will face in the decades ahead" (NGSS, 2013, Appendix I). Further, the Standards for Technological and Engineering Literacy (STEL; (International Technology and Engineering Educators Association [ITEEA], 2020) explicitly call for the integration of engineering practices within other content areas. The STEL were developed based on the 21st Century skills and the Engineering Habits of Mind, which include systems thinking, problem-finding, visualizing, improving, creative problem-solving and adapting. In the science classroom, children are increasingly being asked to define problems, design solutions, develop models, apply computational thinking, collaborate with others, and communicate solutions.

While the terms science, STEM and engineering tend to be used interchangeably in the field of education, STEM education is defined as the integration of the knowledge and skills of two or more STEM disciplines applied to a real-world problem (English, 2016). Similarly, engineering education focuses on developing the engineering design process (EDP) skills within the science curriculum (and increasingly among other disciplines including math and art). The ITAAE (2020) more specifically delineates the difference as an emphasis on "technological products, design, and technology/society interactions…and the use of engineering design and application of engineering habits of mind…" (p. 5). While several models of the EDP exists for different age groups, the model generally includes determining the problem to be addressed, defining the constraints and requirements, researching and brainstorming ideas, planning (and drawing models), building, then testing, and then working with other "engineers" to improve the design. In addition to being an integral part of engineering education, the EDP also supports student development of problem-solving skills (NSTA, 2011). Given that the NGSS asks students to demonstrate the science and engineering practices, this article focuses on the EDP which most closely aligns with those practices.

Self-efficacy has long been linked to achievement and other desirable outcomes (Larson et al., 2015). Assessing self-efficacy so that it can be improved through curricular interventions is necessary to improving student outcomes in engineering and STEM education. While there are scales for adults to measure engineering self-efficacy and for children there are STEM self-efficacy measures, presently there are no assessments of engineering self-efficacy for children. To address this gap in measurement, this article reports the development and validation of the Children's Engineering Self-Efficacy Scale (CESES). This scale is based on the elements of the EDP and assesses the degree to which students are confident in their abilities to engage with all steps of the EDP including planning, problem solving, improving solutions, and working with others to improve a design. This article also describes differences among various subgroups on the CESES.

The Role of Self-efficacy in Engineering Education

People develop beliefs about their own self-efficacy by interpreting and synthesizing multiple sources of information about their abilities (Pfitzner-Eden, 2016). Bandura's (1997) sources of information self-efficacy include mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states, although mastery experiences have the most significant contribution to self-efficacy.

As noted previously, self-efficacy has been correlated with improved educational outcomes including graduation rates, career choice, engagement, and effort (Larson et al., 2015; Ponton et al., 2001; Webb-Williams, 2018). Science self-efficacy has been examined for its impact among groups including teachers and girls. Previous research showed that higher self-efficacy for science teaching demonstrated increased time spent teaching science, responsiveness to students, motivation to teach more creatively and to try new methods (Utley et al., 2019). Research on pre-service teachers has also revealed that exposure to engineering design can significantly improve engineering pedagogical content knowledge and self-efficacy (Perkins Coppola, 2019; Yesilhurt et al., 2021).

With respect to STEM education, Bandura et al (2012) examined the role of self-efficacy beliefs on children's career paths among over 270 children. Children's perception of their efficacy and academic self-efficacy, rather than their academic achievement, was a major predictor of perceived occupational self-efficacy, in other words, their ability to do well in particular careers. Importantly, gender differences showed that male students had higher math self-efficacy and female students had higher language self-efficacy. Given the relationship between self-efficacy and career choices, these results have major implications for steering women into STEM based fields. Other researchers have also linked self-efficacy to career participation (see for example, Caspi et al, 2019; Caspi et al, 2020; Conradty et al., 2020).

Research from the Boston Children's Museum, developers of the Engineering is Elementary program, has shown the positive impact of engineering curriculum on children's science learning and problem-solving skills (Cunningham et al., 2019). Research on children's self-efficacy in STEM or science has examined children's perceived confidence in science and math and how it impacts constructs such as career choice and performance (Bandura et al, 2001; Betz & Hackett, 2000). A great deal of research on teacher's self-efficacy in STEM and engineering has been produced. And it is certainly the case that teacher self-efficacy impacts students' performance. For example, a meta-analysis conducted by Menon et al. (2024) showed that that gains in teacher self-efficacy in STEM and engineering are limited. However, no research has been conducted on children's self-efficacy in engineering design. Given the relationship between self-efficacy, achievement, and other outcomes assessing and intervening to improve, engineering design self-efficacy is a necessary element to facilitating the success of engineering design curriculum.

Existing Scales of Self-Efficacy and Engineering

Much research on self-efficacy and engineering examines the impact of teacher self-efficacy on student efficacy because teacher efficacy can predict teaching behaviors. For example, Yoon et al. (2014) developed a scale to assess teacher self-efficacy to teach engineering design. Items on the scale primarily dealt with teacher confidence to teach engineering with statement such as "When students do better than usual in engineering, it is often because I exerted a little extra effort" (p. 484). Hammack and Ivey (2017) also examined teacher engineering self-efficacy (ESE) and found that they largely had low engineering design self-efficacy in addition to having low self-efficacy for pedagogical content knowledge related to engineering design. These results are unsurprising and point to the need for significant efforts to offer professional development to promote knowledge and confidence to teach children engineering.

Self-efficacy scales developed for engineering students are limited to adolescents and adults. Amato-Henderson et al. (2007) developed the Longitudinal Assessment of Engineering Self-Efficacy for middle and high school students. The scale contains statements about students' perceptions of success in their engineering classes and those around them as well as their opportunities for advancement in engineering in the future. The authors found that while the engineering event around which the study was based did not improve self-efficacy in engineering, knowing an engineer did significantly contribute to positive self-efficacy. This drives home the value of having access to a role model who is participating in the STEM workforce.

Carberry et al. (2010) developed the engineering self-efficacy scale for adults ages 20-62. Like the present study, they sought to evaluate ESE by examining participant's confidence and motivation to participate in each step of the EDP including identifying a design need, developing solutions, selecting the best design, constructing a prototype, testing and communicating a design, and then redesigning (improving). They found that self-efficacy in engineering was highly correlated to outcome expectancy and motivation and negatively correlated to anxiety.

While the above is sampling of available engineering self-efficacy scales, they are limited in utility for the present study for several reasons. Carberry's (2010) scale was designed for adults but used language most like the scale developed for the current study in that participants are asked to rate their confidence in doing specific engineering design tasks, such as formulating questions. Yoon et al (2020) explicitly ask teachers to rate their confidence to teach engineering as opposed to the specific engineering design tasks. One notable gap in the literature is the lack of assessment as to whether participant's self-efficacy is correlated to their actual knowledge or abilities. While this later concept is not explored in the current study, it emerges as a method of convergent validity.

STEM Self-Efficacy

Although scales for measuring children's self-efficacy for participating in the EDP are lacking, several scales have been developed to measure STEM self-efficacy. The Activation Lab (2019) has developed a variety of measures to assess children's and adolescent's affective domains as they relate to science and STEM. Most salient to the current study is the competency beliefs in STEM survey, which provides statements asking participants to rate their level of agreement about their abilities to read science texts, solve math problems and engage in technology. These surveys were developed for students in grades 5 and above and do not address the EDP as discussed further in the methods section.

Luo et al. (2021) recently published findings on their STEM self-efficacy scale developed for 844 primary school children in grades 4-6. The scale was developed to assess self-efficacy to complete STEM activities. Many of the items on the scale are closely aligned with the EDP; examples include "I am able to define the problem to be solved...I am able to design solutions to the problem." (p. 414). The scale had a CA of .90 with a single factor structure. Analysis of their findings showed that participating in school and out-of-school STEM activities significantly predicted SE in STEM, while age and gender did not.

This review of research makes it clear that there is an established need to assess ESE among children who are now being introduced to engineering design process skills via the NGSS. No published studies have described a scale suitable use for children that assesses self-efficacy in using the engineering design process. If children and youth will be expected to demonstrate these skills, evaluating their ability, knowledge of, and self-efficacy is integral to assessing curricular effectiveness and other interventions. Accordingly, the following section describes the development and validation of the CESES.

Methods

Participants

Participants were students in a variety of educational settings in which an engineering design program was implemented. A total of 263 surveys were collected as some participants took both a pre and post-test. A total of 212 of those surveys, which were only post-tests, were used for the reliability and validity evaluation. Of those 212 participants, 144 were girls and 68 were boys. There were more girls than boys because 81 participants were in grades 3-8 attending a four-week Summer STEM program for girls.

Participants also included 131 students who were taught four-week long engineering design units during the regular school day. Twenty-one of these students were from a gifted school. Four children were designated as students with learning disabilities. Most students were in 6th grade (n = 130), 24 were in 3rd, 9 were in 4th, 33 were in 5th and 16 were in 7th grade.

All students were taught engineering design units in which the culminating project was engineering design challenge. All students were taught the engineering design process through the unit. Children in the summer program were taught two short units and children in school were taught one longer unit; however, all instruction was four weeks long. Most participants attended public schools in Nassau County, Long Island in New York State. However, one small group of students were 5th graders in a private gifted and talented school. Two instructors taught in the summer program, and six taught in the regular school program. All instructors were graduates of a master's degree program in elementary STEM education and were certified elementary school teachers. These instructors took courses on children's engineering, STEM curriculum design and produced master's thesis implementing engineering curriculum.

While some groups of students received the questionnaire as both a pre and post-test to an engineering curriculum intervention, only post-test data were used for the reliability tests and factor analysis because not all children took both a pre and post-test. Because the scale was administered in different scenarios, via a summer program, a regular school setting, and in a gifted school, some children were not present for the administration of the posttest. The researcher decided to use the post-test rather than pre-test because all students would have been exposed to engineering design vocabulary and concepts. A limitation of this approach is that all students received training in engineering design and were likely more self-efficacious than children who did not receive such training, as is further explained below regarding the development of the scale. Two hundred and twelve students took a post-test. Post test data were also used to compare performance on the CESES among genders, school type (summer vs. regular), and grade levels.

Development and content validation of the CESES

The development of the initial scale began with a review of published scales on children's self-efficacy in STEM, science and math. In addition, existing scales on ESE for older participants were examined. The Engineering Design Self-Efficacy Instrument developed by Carberry et al (2010) was designed for adults and was validated with participants ages 20-62. The questionnaire includes 17 statements, three of which are negatively phrased. The items were based on the eight-step model of the EDP developed by the Massachusetts Department of Education (2001/2006). Participants in Carberry's study were asked to rate their level of anxiety, confidence, and motivation to do each listed engineering task (e.g., research a design need, communicate a design, and redesign). While Carberry et al.'s scale was conceptually closest to the CESES, other scales written for children provided a more relevant framework for the development of the items.

Accordingly, the STEM competency beliefs scale by Activation lab (2019) provided the basis for the structure of the statements. The statements themselves were designed to follow the engineering design process as presented in the Engineering is Elementary curriculum (EIE.org, 2020). The items were designed to assess the following phases of the EDP: ask (asking questions), imagine (brainstorm), plan (with a detailed sketch), create (build) and improve (test, consult with other engineers, and improve the design, and test again). Because the EDP is designed to be conducted collaboratively, several items included the phrase "…when working with others…"

Bandura's (1997) sources of self-efficacy served as the conceptual basis for developing the items. Because mastery experiences are the biggest contributors to self-efficacy, the phrasing of the items focused on a participants' confidence in engaging in a part of the engineering design process. Most of the items in the CESES begin with the phrase "When given a design challenge, I can…" The survey is divided into two parts to allow students space to pause and think between two groups of items. The statements fall into the following categories that represent the EDP as described above: confidence in solving problems, knowing which questions to ask, knowing when and how to draw a plan of their ideas, working with others, knowing that failing and improving is part of the process, knowing how to apply constraints and requirements. Two negatively phrased items were included in this scale. Negatively phrased items are included in scales to mitigate response bias, particularly salient when assessing children as they are more likely to acquiesce with positively phrased statements for social acceptability reasons. The inclusion of the negatively phrased items allowed participants to consider the concept being assessed through a rephrasing. In addition, negatively phrased items are often included in scale development to inhibit participants from reading the statements too quickly and responding positively to each item.

The first version of the questionnaire included a scenario about being asked to develop a solution for designing a roller coaster for an amusement park. The scenario was then followed by questions intended to determine whether students were able to do the task for which they reported self-efficacy. However, this section made the scale difficult to score as the questions needed to be open-ended. Future revisions or predictive validity assessments might include this type of scenario.

Because the scale is intended to assess self-efficacy of specific engineering skills, children may not be familiar with many of the vocabulary words until they have received instruction on the engineering design process and/or completed at least one challenge. It is recommended that for pre-post testing, that the pre-test be given after one brief challenge or after students have been taught the EDP. Students should then be post-tested after the established curriculum is presented to measure growth. Pre-testing prior to any instruction may fail to capture whether a child has low self-efficacy or does not know the vocabulary. In this study, all students had some experience with engineering design prior to being given the pre-test CESES. Post-test data only was used for the validation study.

Content validation

Content validity was assessed by sending the original version of the scale to an Engineering Department faculty member who is an expert in children's engineering design and teaches courses on the design of engineering curriculum for children. This expert also recommended the inclusion of the scenario but agreed that its inclusion took away from the intent of the scale to measure SE only. This faculty member also recommended revisions to wording and suggested reviewing the EDP developed by the Massachusetts Department of Education (2001/2006). Additionally, another scientist working on the engineering programs Wise Guys and Girls and STEMgineering reviewed the scale. Finally, the scale was reviewed by five elementary school teachers who had graduated from a STEM Master's degree program, who had all written engineering design curriculum and conducted a related study for their theses. Revisions to wording was made based on this feedback. The items from resulting from these efforts appear in the results section of this article.

Data collection procedures

The survey was administered before and after the following STEM interventions. The first STEM intervention was a summer STEM program for girls in grades 3-7. The program was four weeks long, two days a week for half days. Students engaged in two engineering design units with design challenges including electrical engineering, biotechnology engineering, designing race cars, and designing prosthetics for animals. All units involved learning science content needed to address the engineering design challenge. The last two lessons of each unit focused on planning and designing a solution to the challenge, then testing, improving and receiving feedback from other "engineers". The study was conducted over two years, so each year participants engaged in different design challenges. For girls who participated both years, only their post-test (2nd) survey was used in this study. Students were read the survey allowed and instructors noted which students had difficulty completing the survey.

For the 131 students who completed the CESES during in-school programs, all were exposed to instruction on the EDP and had experience completing small engineering design challenges such as building towers using index cards prior to taking the survey. All students were then taught an 8-lesson science unit that culminated in an engineering design challenge. Unit topics included designing an ecosystem for a particular animal, designing a new planet under specific conditions, designing a constellation, and designing housing to withstand extreme weather.

Results

Reliability

To assess the psychometric properties of the CESES, 212 post-tests were analyzed. To determine the internal consistency and reliability of the scale, a Cronbach's alpha .81 was derived for the 17 items scale, indicating good reliability (Hair et al., 1998). Alphas for all items ranged from .791 to .826.

Construct validity

A principal components analysis with Varimax rotation was conducted to determine construct validity and variance explained. Varimax rotation was employed as it maximizes variance and is useful in analyzing multiple latent variables that are not correlated (Dilbeck, 2017). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .841 and Bartlett's Test of Sphericity was significant at $p < .001$ with a value of 840.812; both measures indicated that the data were suitable for factor analysis.

Table 1.

Rotated component matrix and factor loadings

	Factor				
	1	2	3	4	5
Q9	.724	.066	-.042	.057	-.142
Q12	.653	.133	.064	.091	.167
Q11	.619	.364	-.009	.070	.175
Q1	.616	-.006	.237	.023	-.192
Q13	.600	.178	.052	.120	.247
Q2	.581	.117	.104	.204	.013
Q8	.515	-.182	.283	.404	-.021
Q7	.453	.386	.434	.162	-.079
Q10	.432	.246	.366	.170	-.402
Q4	.198	.799	-.082	.146	.034
Q5	.129	.774	.247	.177	-.068
Q15	.001	.181	.694	-.172	.269
Q6	.186	-.065	.693	.234	.139
Q3	.121	.163	.118	.748	.100
Q17	.186	.202	-.047	.732	-.033
Q14	.218	.011	.080	-.039	.736
Q16	-.106	-.011	.224	.131	.628

Five factors with an Eigen value greater than 1.0 emerged from these analyses. This five-factor model represented a combined variance of 56.72%. All but two items loaded on distinct factors at loadings greater than .454, with most factors loading at .6 and above. Items 7 and 10, loaded between .4 and .5 and loaded on several other components at the .2-= to .3 range. Both questions referred to improving a design and conceptually belong to the factor one which articulates several steps of the EDP. Removing these two items resulted in a four-factor model, in which the variance explained decreased to 51.81 and reliability decreased to .77. To maintain reliability, the 5-factor model was retained (Table 1). See Table 2 for the list of items, with loadings for each factor.

Table 2.
Five factor model with items statements arranged by factor

Factor	Eigenvalue	Total variance	Item number	Item
1.Problem solving and asking questions	4.67	27.468	1	I am good at solving problems
			2	When given a design challenge, I know which questions to ask in order to get started
		Low loading	7	If my design doesn't work well the first time, I can improve it and test it again
			8	I can ask other classmates for their advice on my design
			9	I am good at coming up with new ideas to solve problems
		Low loading	10	I can come up with more than one solution and pick the one that works the best.
			11	When given an engineering design challenge, I can figure out what needs to be included in the solution
			12	When given an engineering design challenge, I can create a solution using only materials given to me
			13	When given an engineering design challenge, I can create a solution in the amount of time I'm given by my teacher
2.Modeling and planning	1.54	9.069	4	When given a design challenge, I can create a model of my solution to start solving the problem
			5	When given a design challenge, I can draw a plan to solve the problem
3. Working with others	1.28	7.524	6	When given a design challenge, I am good at working with the other people in my group to come up with a solution
			15	I don't think it's helpful to ask other classmates for their feedback

Factor	Eigenvalue	Total variance	Item number	Item
4.Brainstorming	1.04	6.134	3	When given a design challenge, I first think about how I can solve the problem before I start building
			17	When I'm given a design challenge, I do some research first
5.Lack of improving and planning	1.03	6.039	14	When my design doesn't work the first time, I don't feel like finishing the project
			16	When I'm given a design challenge, I start building right away

EDP Self-Efficacy Scores

To examine differences between groups, the total sample of 212 post-tests were used. The mean score for all participants was 53.8 and the standard deviation was 6.94. The range of scores was between 33-68 (Table 3). To allow researchers to determine the strength of self-efficacy among children, cut off scores were derived using standard deviations. The minimum possible score was 17 and the maximum score was 68. Since the mean score was 53.7, the average scale rating for each item was 3 points, indicating that all participants had reasonable self-efficacy after the post test. These cutoff scores are only a guideline and based on this sample's performance. Other researchers may wish to use the raw scores for other types of analyses. Because many of the participants in this study had been exposed to engineering curriculum previously and were taught by trained STEM teachers, the scores reported in this study may be higher than typical.

Table 3.
Cutoff scores for ESE

Score range	Category	Number of participants scoring	Percent
61-68	high ESE	43	20.3
53-60	average to good ESE	86	40.4
45-52	average to low ESE	58	27.2
17-44	low ESE	25	11.8

These score ranges demonstrate that most participants were average/good too high in ESE. These higher scores make sense in the context of the students' exposure to other STEM interventions including the summer program and the fact that all instructors in the study were trained in engineering design instruction.

Differences between genders on engineering self-efficacy

Several Analysis of Variance (ANOVAs) tests were performed to determine whether there were significant differences between various groups on the CESES. In the use of an ANOVA there is an assumption that variances among groups within a population are equal. Because subgroups within the sample varied, a Levene's test for homogeneity of variance was conducted. The Levene's test was not significant at p = .795 indicating that the assumption of homogeneity of variance has been met. Therefore, the F tests are interpretable.

Girls demonstrated a mean score of 54.55 (n = 144) and boys 52.07 (n = 68). There was a significant difference between girls and boys (F (1, 2120 f = 5.950, p < .01). However, there was an uneven representation of genders in this study. Many of the participants in this study were girls, and many of those girls were attendees at the summer STEM program which focuses on engineering. Many of the girls in the summer program had been attending for 2-4 years.

Differences between grade levels on engineering self-efficacy

Differences between grade levels were also assessed given the wide distribution of grades. There was a significant difference between grades on ESE, (F(4, 212) f = 2.753, p < .05) 5th graders (n = 33) had the highest mean score at 56.69, while 3rd graders had the lowest mean score at 51.62 (n= 24). See Table 4 for means. Post hoc analyses using a Tukey test revealed that the significant difference was between grades 3 and 5, p < .05. There was trending difference between grades 6 and 7, p = .064

Differences between summer program participants and school programs

Finally, means between students in the summer STEM program for girls and students who took the CESES during regular school programming were analyzed. Participants in the summer program scored a mean of 54.24, while those in regular school scored a mean of 53.45. These differences were not statistically significant, p = .426.

Table 4.
Comparison between grades on CESES

Grade	**Mean**	**N**
3	51.6250	24
4	53.8889	9
5	56.6970	33
6	53.1308	130
7	55.9375	16
Total	**53.7594**	**212**

Discussion

This study describes the development and validation of an ESE scale developed for children and young adolescents. The results demonstrate that the CESES is a valid and reliable tool that may be used to assess the self-efficacy of children and youths in grades 3-7 after participating in an engineering design unit. While the results of the factor analysis showed that the items from the five-factor model loaded onto distinct factors, these factors did not cleanly emerge from the EDP –demonstrating that the EDP skills are linked in some cases. In fact, most items in factor one related to several steps in the EDP including problem solving, asking questions and improving indicating the design process is conceptually one latent variable. The other factors were more related to the mechanics of the EDP such as impulsive decision making and working with others. For example, working collaboratively emerged as its own factor. Working collaboratively can be characteristic of all the EDPs but only emerged as its own factor in questions asking specifically about working with others. The other factors that emerged related to lack of planning and improving (impulsivity), problem solving and questioning in general, working with others, and planning and modeling. These factors make sense in the context of the science and engineering practice described in the NGSS standards. A self-efficacy scale to complete STEM activities, with items covering some aspects of the EDP, developed by Luo et al (2021) found a one factor solution for a 10-item scale.

Mastery experiences strongly predict self-efficacy (Pfitzner-Eden, 2016). Examination of the average scores among subgroups showed that participants had good to high scores. Most participants were girls who were the participants in the summer program in which they engaged in several engineering design projects. These results affirm that frequent exposure to engineering concepts, particularly improvement as an objective in the process, will increase confidence in one's abilities.

In this study most participants had average to high average scores on the CESES as indicated by the cutoff scores derived from this sample. The participants in this sample were all taught by relatively new teachers with 1-3 years of experience but trained in the design and use of engineering design challenges. All instructors graduated from an elementary education STEM Master's program in which engineering is the focus. Further, all instructors completed theses that assessed the effectiveness of a STEM (engineering design) unit. While post-test scores were used for this study, nearly all participants had been exposed to an engineering design challenge previously and so their scores may be higher than those of other children. As is a risk in many studies that include self-report, particularly among children, high scores may be a result of social desirability. This risk is another argument for validating the results with an assessment of true engineering ability and knowledge.

Results also showed no significant differences between males and females on self-efficacy; however, many of the participants were female. These findings support those of Luo et al (2021) who did not find differences between genders on self-efficacy, but did find that participation in both in-school and out-of-school STEM experiences predicated self-efficacy in STEM. In this study no differences were found between participants in the summer program and those who were exposed to the engineering units during the school day, but what all participants had in common was exposure to at least one engineering design challenge and unit, and thus familiarity with the vocabulary.

Researchers and educators who choose to assess CESES with children who have no experience with engineering design will very likely find much lower scores. Because the vocabulary used in the CESES is discipline-specific, it is important that those administering the scale take time to expose students to the vocabulary through readings or other related activities. To preserve the validity of the scale when it is being used a pre-test, students should not have had the experience of completing and engineering design challenge previously, or at least that experience should be statistically controlled for. In addition, for true growth in content knowledge and engineering design ability and self-efficacy, meaningful instruction should focus on design challenges that solve a relevant and local problem of importance to children. Repeated exposure to engineering design challenges through a variety of disciplines will not only improve engineering self-efficacy, but also problem-solving skills.

Limitations and Recommendations for Future Research

The limitations of this study include the diversity of age and grade ranges. All grade ranges scored in the average to high realm, but there were differences among 3, 5 and 6th graders. The large number of 6th graders also may have skewed scores upwards. Future research may target a smaller age band for which this scale is appropriate. Third graders scored the lowest on this scale; this may have been a function of literacy level or less exposure to engineering design challenges. As survey taking is largely a literacy endeavor, those participants who were emergently literate may have difficulty completing the survey. The few students in this study with IEPs were read the survey aloud. Future research should include evaluating the readability of the survey for younger participants.

The study included new teachers trained in engineering design. Future research might include more experienced teachers with less engineering design training. The true usefulness of the CESES will show a baseline of where students stand in the ESE prior to introduction to the engineering curriculum. It will also show the impact of teachers with and without engineering training and professional development. Finally, the validation study only included 212 participants in one suburban area of New York State. Future research should assess the validity of this scale on children in other regions of the United States as well as internationally.

Future research should also examine predictive and concurrent validity. Correlating the scores of the CESES against other science or STEM self-efficacy measures would better support the validity of this measure. The development of the CESES allows researchers to use the scale to assess the effectiveness of engineering curriculum and other programmatic interventions on students' ESE. Accordingly, this scale may be used as a pre and posttest in various contexts. Lastly, future research should include assessing the match between students reported self-efficacy and their actual ability as intended by the scenario included in the initial version of the scale.

Note: Please write to the author for a copy of the CESES or to obtain permission to use the scale.

References

Activation lab. (2019). *STEM competency beliefs scale*. ActApp: The Activation Lab Evaluation Toolkit. https://activationlab.org/toolkit/

Amato-Henderson, S., Mariano, J., Cattelino, P., & Hannon, B. (2007). Who You Know Does Matter in Engineering Self Efficacy. In *Proceedings of the 2007 ASEE North Midwest Sectional Conference*.

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

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child development*, *72*(1), 187-206. https://doi.org/10.1111/1467-8624.00273

Betz, N. E., & Schifano, R. S. (2000). Evaluation of an intervention to increase realistic self-efficacy and interests in college women. *Journal of Vocational Behavior*, *56*, 35–52. https://doi.org/10.1006/jvbe.1999.1690

Carberry, A. R., Lee, H. S., & Ohland, M. W. (2010). Measuring engineering design self-efficacy. *Journal of Engineering Education*, *99*(1), 71-79. https://doi.org/10.1002/j.2168-9830.2010.tb01043.x

Caspi, A., Gorsky, P., Nitzani-Hendel, R., Zacharia, Z., Rosenfeld, S., Berman, S., & Shildhouse, B. (2019). Ninth-grade students' perceptions of the factors that led them to major in high school science, technology, engineering, and mathematics disciplines. *Science Education*, *103*(5), 1176–1205. https://doi.org/10.1002/sce.21524

Caspi, A., Gorsky, P., Nitzani-Hendel, R., Zacharia, Z. C., Rosenfeld, S., Berman, S., & Shildhouse, B. (2020). Children's perceptions of the factors that led to their enrolment in advanced, middle-school science programmes. *International Journal of Science Education*, *42*(11), 1915-1939. https://doi.org/10.1080/09500693.2020.1802083

Conradty, C., Sotiriou, S. A., & Bogner, F. X. (2020). How creativity in STEAM modules intervenes with self-efficacy and motivation. *Education Sciences*, *10*(3), 1-15. https://doi.org/10.3390/educsci10030070

Cunningham, C. M., Lachapelle, C. P., Brennan, R. T., Kelly, G. J., Tunis, C. S., & Gentry, C. A. (2019). The impact of engineering curriculum design principles on elementary students' engineering and science learning. *Journal of Research in Science Teaching*, *57*(3), 423-453. https://doi.org/10.1002/tea.21601

Dilbeck, K. (2017). Factor analysis: varimax rotation. In *The SAGE Encyclopedia of Communication Research Methods* (Vol. 4, pp. 532-533). SAGE Publications. https://doi.org/10.4135/9781483381411.n191

English, L. (2016). STEM education K-12: perspectives on integration. *International Journal of STEM Education*, *3*(3), 1–8. https://doi.org/10.1186/s40594-016-0036-1

Hair, J.F., Tatham, R.L., Anderson, R.E., & Black, W. (1998). *Multivariate data analysis*. Prentice-Hall

Hammack, R., & Ivey, T. (2017). 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

International Technology and Engineering Educators Association. (2020*). Standards for technological and engineering literacy: The role of technology and engineering in STEM education*. www.iteea.org/STEL.aspx

Larson, L. C., & Miller, T. N. (2011). 21st century skills: Prepare students for the future. *Kappa Delta Pi Record*, *47*(3), 121-123. https://doi.org/10.1080/00228958.2011.10516575

Larson, L. M., Pesch, K. M., Surapaneni, S., Bonitz, V. S., Wu, T. F., & Werbel, J. D. (2015). Predicting graduation: The role of mathematics/science self-efficacy. *Journal of Career Assessment*, *23*(3), 399-409. https://doi.org/10.1177/1069072714547322

Luo, T., So, W. W. M., Li, W. C., & Yao, J. (2021). The development and validation of a survey for evaluating primary students' self-efficacy in STEM activities. *Journal of Science Education and Technology*, *30*(3), 408-419. https://doi.org/10.1007/s10956-020-09882-0

Massachusetts Department of Education (2001/2006). *Massachusetts science and technology/engineering curriculum framework*. Massachusetts Department of Education.

Menon, D., Wieselmann, J. R., Haines, S., & Asim, S. (2024). A meta-synthesis of the literature on science & engineering teaching self-efficacy: current gaps and future research directions. *Journal of Science Teacher Education*, 1-24. https://doi.org/10.1080/1046560X.2023.2297499

NSTA. (2011). *Quality Science Education and 21st-Century Skills*. https://www.nsta.org/nstas-official-positions/quality-science-education-and-21st-century-skills

Perkins Coppola, M. (2019). Preparing preservice elementary teachers to teach engineering: Impact on self-efficacy and outcome expectancy. *School Science and Mathematics*, *119*(3), 161-170. https://doi.org/10.1111/ssm.12327

Pfitzner-Eden, F. (2016). Why do I feel more confident? Bandura's sources predict preservice teachers' latent changes in teacher self-efficacy. *Frontiers in Psychology*, 7, 1-16. https://doi.org/10.3389/fpsyg.2016.01486

Ponton, M. K., Edmister, J. H., Ukeiley, L. S., & Seiner, J. M. (2001). Understanding the role of self-efficacy in engineering education. *Journal of Engineering Education*, *90*(2), 247-251. https://doi.org/10.1002/j.2168-9830.2001.tb00599.x

Utley, J., Ivey, T., Hammack, R., & High, K. (2019). Enhancing engineering education in the elementary school. *School Science and Mathematics*, *119*(4), 203–212. https://doi.org/10.1111/ssm.12332

Webb-Williams, J. (2018). Science self-efficacy in the primary classroom: Using mixed methods to investigate sources of self-efficacy. *Research in Science Education*, *48*(5), 939-961. https://doi.org/10.1007/s11165-016-9592-0

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), 1-15. 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