Development and Initial Validation of the Student Interest and Choice in STEM (SIC-STEM) Survey 2.0 Instrument for Assessment of the Social Cognitive Career Theory Constructs

Sarah A. Roller, Sandra A. Lampley, Monica L. Dillihunt, Michael P. J. Benfield, Sampson E. Gholston, Matthew W. Turner, et al.

⌂ Springer

Springer

Development and Initial Validation of the Student Interest and Choice in STEM (SIC-STEM) Survey 2.0 Instrument for Assessment of the Social Cognitive Career Theory Constructs

Sarah A. Roller[1] · Sandra A. Lampley[1] · Monica L. Dillihunt[1] · Michael P. J. Benfield[2] · Sampson E. Gholston[3] · Matthew W. Turner[2] · Andrew M. Davis[2]

Abstract

In response to a growing need for STEM professionals, this study reports the results of the initial validation of a refined survey instrument that purports to measure the five constructs of the social cognitive career theory framework within the subjects of mathematics, science, engineering and technology. To investigate the instrument's reliability and psychometric properties, we administered this 45-item survey to students in grades 4–12. Reliability and validity (content and construct) were assessed using Cronbach's coefficient alphas and structural equation modeling. Path coefficients for the five constructs indicated weak to moderate influences on the subscales, and the goodness-of-fit indices demonstrated that the model is acceptable. Initial results indicate the survey has the potential to collect reliable and valid data and suggest the instrument may be helpful in measuring students' interests and choices in STEM careers for research, partnerships, and curricular development. Additionally, results highlight two areas for further investigation, which include the impact conscientious or random responders have on the survey's psychometric properties and what constitutes satisfactory Cronbach's alpha for results to be interpreted in a significant way.

Keywords STEM education · Instrument design · Instrument validation · Social cognitive career theory

Introduction

In this age of rising global competition, proficiency in science, technology, engineering, and mathematics (STEM) disciplines is viewed as being increasingly essential in today's workforce. STEM workers propel innovation and competitiveness by generating new ideas, new companies, and new industries (Noonan 2017). Globally, employers are faced with the task of recruiting more individuals into STEM industries (Marginson et al. 2013). Many countries are not producing the amount of STEM professionals that are needed in order to meet the workforce demands (Wilson et al. 2016). In the USA, STEM occupations have increased by 14% since 2008, while other occupations have grown only 1.7% (Noonan 2017). It is projected that the USA will need approximately one million more STEM professionals than it will produce over the next decade to meet this growing demand (The President's Council of Advisors on Science and Technology; PCAST 2012). To accomplish this goal, a 34% annual increase in the number of students who receive undergraduate STEM degrees is needed (PCAST 2012). Despite this need, only 24% of incoming college students declare a STEM major (Shapiro and Sax 2011), and less than half of the students who declare a STEM major graduate with a STEM degree (Price 2010). Policymakers, researchers, and educators are emphasizing the importance of educating students for STEM-related jobs, but too few of US high school graduates are ready for college coursework or careers in STEM areas (American College Testing; ACT 2017).

In order to understand why many countries are not producing the workers in STEM fields at the rate needed, it is important to understand the factors that motivate students to pursue or not to pursue a career in STEM fields. While much of the focus has been on increasing student achievement in science

✉ Sarah A. Roller
sarah.roller@uah.edu

1 Department of Curriculum and Instruction, The University of Alabama in Huntsville, Huntsville, AL, USA

2 Systems Management and Production Center, The University of Alabama in Huntsville, Huntsville, AL, USA

3 Department of Industrial and Systems Engineering and Engineering Management, The University of Alabama in Huntsville, Huntsville, AL, USA

and mathematics (National Research Council 2011), it is important to understand other factors involved, such as student interest and student choice in STEM. Considering this need for information, this article describes the second phase of development of a survey designed to measure students' interest and choice in STEM with populations ranging from 4th–12th grade.

Theoretical Framework

As part of a National Science Foundation (NSF) grant funded project to investigate the challenges and barriers for students to pursue a STEM career, the researchers decided to utilize *social cognitive career theory* (SCCT) as a theoretical framework to understand how individuals determine academic and career choices (Lent and Brown 2006, 2008; Lent et al. 1994, 2000). Derived from Bandura's social cognitive theory (1986), the SCCT framework explores the way people's behavior and environments can influence one another (Lent 2013). The aim of SCCT is to explain four interrelated aspects of career development: (a) how basic academic and career interests develop, (b) how educational and career choices are made, (c) how academic and career success is obtained, and (d) how satisfaction or well-being in the work environment is experienced (Lent 2013). Correspondingly, the SCCT framework explores four interrelated models of career development including interest, choice, performance (Lent et al. 1994), and satisfaction (Lent 2013). Within each of the four models of SCCT, there are three sociocognitive mechanisms that are believed to apply important influences to career development. The sociocognitive mechanisms include self-efficacy beliefs, outcome expectations, and personal goals (Lent et al. 1994). The three mechanisms will be further described in the subsequent paragraphs.

Self-efficacy is the belief that an individual has pertaining to his/her ability to perform a certain behavior or task (Bandura 1997). Self-efficacy beliefs are dynamic and are specific to a particular activity. It is thought that self-efficacy is derived from four primary sources of information: experiences of mastery, vicarious experiences (e.g., observing similar others), social persuasion, and emotional and physiological states (Bandura 1997). SCCT assumes that individuals are likely to develop interest, choose to pursue the interest, and therefore perform better at tasks and/or activities in which they have a stronger self-efficacy (Lent 2013).

Outcome expectations are personal beliefs about the result or consequences of performing particular behaviors (Lent 2005; Lent 2013). Self-efficacy is believed to affect outcome expectations (Lent 2013), and both can influence whether an individual will choose to persist or avoid a particular activity (Bandura 1986; Lent et al. 1994; Lent 2013). For example, an individual is more likely to participate in an activity if she or he views the engagement will lead to a valued or positive outcome (e.g., social and self-approval, tangible rewards, desirable work conditions).

Bandura (1986) describes personal goals as an individual's intent to engage in a particular activity or perform a particular outcome. There are two types of personal goals described in the SCCT framework: they include choice goals (the type of activity or career an individual wants to pursue) and performance goals (the quality of performance one plans to accomplish within a given task; Lent 2013). According to the SCCT framework, both choice goals and performance goals are impacted by an individual's self-efficacy and outcome expectations (Lent 2013).

The Interest Model and the Choice Model of the SCCT Framework

Although SCCT includes four distinct models, the researchers decided to build the instrument around the interest and choice models. Both of these models demonstrated affordances to be used broadly, since students could be asked general, selected response questions about their interests and potential choices, whereas the performance model and the satisfaction model would require more student-specific contexts and extended response items. Therefore, the researchers decided to leverage the interest model and choice model to learn more about elementary and secondary students' career decisions.

The interest model explores the ways *self-efficacy* and *outcome-expectations* develop students' *interests*. For example, if an individual is exposed to coding, and has received positive feedback on her ability to code, it is reasonable that she will develop a high self-efficacy for coding. If she views that coding will demonstrate a positive outcome in her future, her high self-efficacy combined with her outcome expectations will expectantly increase her interest in coding. This combination of interest, self-efficacy, and outcome expectations is likely to encourage her involvement in coding activities further, which will increase her likelihood to code more often, which can presumably refine her self-efficacy and outcome expectations (Lent 2013).

Adding on, the choice model considers the ways interest, self-efficacy, and outcome-expectations influence *choice goals*, which then motivate *choice actions*, or the efforts to implement one's goals (Lent 2013). For example, a student who is considering a career in engineering (choice goal) might participate in a robotics competition (choice action). However, he may find he has difficulty completing the required tasks, such as designing and building a robot, and may not enjoy working as part of a team. From this experience, he may also conclude that the design environment and rewards available as a future engineer are less suited for him than he initially thought. These learning experiences may encourage the student to revise his self-efficacy and outcome expectations,

leading to a shift in interest and career goals, which could lead to the choice action of finding a different career path.

Thus, both of these narrative stories illuminate how the SCCT interest model and choice model can be applied to elementary and secondary students' career choices. Specifically, the five SCCT constructs that are involved in these models are interest, self-efficacy, outcome expectations, choice goals, and choice actions. Therefore, these are the five SCCT constructs the researchers decided would guide the instrument design process for this study.

Instruments Utilizing SCCT Constructs

The researchers reviewed previously developed instruments for measuring student outcomes (i.e., content knowledge, reasoning skills, psychosocial attributes, attitudes) and found they were either limited to a single STEM subject (mathematics, Adelson and McCoach 2011; Tapia and Marsh 2004), (science, Germann 1988; Kennedy et al. 2016), (technology, Washington et al. 2016), (engineering, Capobianco 2015) or systematically collected data across all STEM subject areas (Erkut and Marx 2005; Kier et al. 2013; Guzey et al. 2014; Tyler-Wood et al. 2010; Unfried et al. 2015). Out of the surveys that considered all STEM subjects (i.e., science, technology, engineering, and mathematics), there were only two surveys that utilized the SCCT framework: the Student Attitudes toward STEM (S-STEM; Faber et al. 2013; Unfried et al. 2015) and the STEM Career Interest Survey (STEM-CIS; Kier et al. 2013). The S-STEM (Faber et al. 2013; Unfried et al. 2015) is structured to measure student attitudes and interests in STEM, and the STEM-CIS (Kier et al. 2013) is also used to measure student interest in STEM. Thus, the focus of the current surveys was only based on the interest model (containing interest, self-efficacy, and outcome expectations constructs), which excludes the importance of student decision-making regarding choice goals and choice actions that might steer a student toward a certain career path.

Purpose

Because career-related interests are grounded in self-efficacy and outcome expectations which foster career goals, and in turn motivate choice actions, the researchers felt it was imperative to develop a survey that measured the five SCCT constructs driving the interest model and the choice model. Therefore, the authors (Roller et al. 2018) developed the Student Interest and Choice in STEM (SIC-STEM) Survey, which employed SCCT's interest model and choice model. The goal of this study is to determine the validity of the SIC-STEM Survey 2.0. The purpose of this paper is to (a) examine the reliability (content and construct) and validity

evidence of data from the SIC-STEM Survey 2.0 for use in assessing student interest and choices in STEM, (b) discuss survey limitations and areas of future investigation, and (c) offer suggestions for appropriate uses of the survey. Specifically the following research questions guided our evaluation of the instrument:

1. What are the Cronbach's alphas for the subscales of mathematics, science, and engineering and technology for the five SCCT constructs (interest, self-efficacy, outcome expectations, choice goals, and choice actions) and what do these calculations suggest about the *reliability* of the SIC-STEM Survey 2.0?
2. What does the inter-rater reliability of coded items suggest about the *content validity* of the SIC-STEM Survey 2.0?
3. Is the structural equation model of the SIC-STEM Survey 2.0 adequate? What does this model suggest about *construct validity*?

Method

Instrument Design

The creation of the SIC-STEM Survey began in 2016 and started with an investigation of the S-STEM instrument by Faber et al. (2013). After operationalizing Lent's (2013) SCCT construct definitions and coding S-STEM items, a team of four researchers noticed that "the S-STEM items were not evenly distributed across the SCCT constructs and in some cases were lacking entirely" (Authors 2018, para. 13). The researchers developed three items for each SCCT construct for mathematics, science, and engineering and technology, resulting in a 45-item survey. Furthermore, an initial pilot study was conducted with 12 high schools across the USA with a total of 196 out of 210 students completing the survey online.

The SIC-STEM Survey pilot data were subjected to exploratory factor analyses (EFA) using SPSS 9.5 Version 22. Additional examination of the survey revealed five distinct components to be used for advance analysis (Roller et al. 2018). Using 0.7 as an acceptable reliability coefficient (Cronbach 1951; Nunnally 1978); the internal consistency reliability for the instrument, the content specific subscales, and the latent constructs within each of the subscales were measured. Appropriate levels of reliability were established (Roller et al. 2018). Data from the pilot of the SIC-STEM Survey yielded evidence for the validity of the items and reliability of the instrument. In an effort to further improve item validity and instrument reliability, constructs yielding a less than satisfactory Cronbach's alpha, below 0.7 (Nunnally

1978), were investigated and revised. These new items (SIC-STEM Survey 2.0) are included in the current sample.

To make sure the survey was accessible for all students, the readability level of each item was investigated by identifying the Flesch-Kincaid grade level for each item. In cases where the reading level was higher than fifth grade, item language was simplified by shortening the sentence and reducing the number of multisyllabic words. However, four items remained in the range of 5.0–6.6 mainly due to including larger words, such as "engineering." Overall average readability, however, for all 45 items was a 2.0, or second-grade level making the instrument appropriate for the sample population.

Data Collection

The SIC-STEM Survey 2.0 was distributed to students in two different packages: one for upper elementary (4th and 5th grade) students and one for secondary (6th–12th grade) students. The upper elementary package contained the 45-item SIC-STEM Survey 2.0, five questions regarding possible career choices, and six questions about the student. The secondary student package contained all of the items from the upper elementary package plus 32 questions on potential barriers that might be faced by the student in going to college or training school after high school. For the purpose of this study, we will only focus on the 45 SIC-STEM Survey 2.0 items that were administered to both groups.

Sample

Initial validation of the SIC-STEM Survey 2.0 was conducted with data gathered from four school systems in the southeastern region of the USA. As shown in Table 1, a total of 821 students, ranging from 4th–12th grade, completed the survey from the four school systems. The distribution of participants by grade varied from 162 from 5th grade (19.7% of the

Table 1 Sample Grade Distribution

Grade	System A	System B	System C	System D	Total
4	19	34	15	4	72
5	20	109	27	6	162
6	0	24	51	0	75
7	0	6	79	1	86
8	18	0	41	92	151
9	23	19	8	37	87
10	8	0	19	19	46
11	16	0	20	42	78
12	9	3	0	30	42
No Response	7	6	4	5	22
Total	120	201	264	236	821

sample) to 42 from 12th grade (5.1% of the sample). A total of 22 (2.7% of the sample) participants provided no response for their grade.

Moreover, Table 2 lists the demographics from each school system sample. The sample from school system A was primarily White students with a significant population of Hispanic and African American students, while the sample from school systems B, C, and D were predominantly African American students. All school systems included in this analysis are Title 1 as shown by the percentage of free and reduced lunch from Table 2.

Psychometric Properties of Instrument

To ensure the instrument produced valid and reliable data, the psychometric properties of the SIC-STEM Survey 2.0 were established. Three domains describe the quality of an instrument: reliability, validity, and responsiveness (Scholtes et al. 2011; Burton and Mazerolle 2011). The authors' assessed the first two domains, specifically reliability and validity, of the SIC-STEM Survey 2.0 in this initial validation study. Content and construct validity were assessed for the validation part of this analysis.

Reliability

To determine internal consistency, Cronbach's (1951) coefficient alpha for the subscales of mathematics, science, and engineering and technology was calculated for the five SCCT constructs of interest, self-efficacy, outcome expectations, choice goals, and choice actions. Analyses were also performed to examine how each item in the construct supported or weakened the reliability of that specific subscale's construct. Reliability for the mathematics subscale ranged from 0.80 for choice goals to 0.53 for outcome expectations (Table 3). Three of the five constructs (interest, self-efficacy, and outcome expectations) did not yield satisfactory Cronbach's alphas of 0.7 (Nunnally 1978).

However, Cronbach's alpha coefficients were also calculated if an item was deleted (Table 3). Further investigation revealed that all three reversed items (M1, M2, and M13) did

Table 2 Sample Demographics

System	African American	Hispanic	White	Other	Free/reduced lunch
A	18%	18%	56%	8%	48%
B	98%	0%	2%	0%	76%
C	69%	1%	29%	1%	78%
D	99%	0%	0%	1%	78%

Table 3 Mathematics subscale reliability coefficients

Construct	Cronbach's alpha	Cronbach's alpha if item deleted		
Interests	0.64*	**M1** (0.82)	M6 (0.34)	M11 (0.37)
Self-efficacy	0.57*	**M2** (0.70)	M7 (0.24)	M12 (0.45)
Outcome expectations	0.53*	M3 (0.25)	M8 (0.36)	**M13** (0.71)
Choice goals	0.80	M4 (0.70)	M9 (0.76)	M14 (0.72)
Choice actions	0.79	M5 (0.74)	M10 (0.73)	M15 (0.66)

M1 indicates mathematics item 1. Bold denotes reverse-worded. * indicates a Cronbach's alpha below 0.7

not support the reliability of the construct that they were intended to support, suggesting that either the wording of the item needed to be altered or that the students may not have noticed the reversed wording. If the reversed items were removed, the remaining items would yield satisfactory Cronbach's alphas for the interest (0.82), self-efficacy (0.70), and outcome expectations (0.71) constructs. Because of the results of this initial reliability analysis, further investigation is warranted to determine if the students were behaving as conscientious or random responders (Marjanovic et al. 2015).

Similar to the mathematics subscale, three of the five constructs (interest, self-efficacy, and outcome expectations) within the science subscale did not have satisfactory Cronbach's alphas, while the choice goals and choice actions constructs were satisfactory (Table 4). However, unlike the mathematics subscale, if the science reverse-worded items were removed, it only provided a satisfactory Cronbach's alpha for the self-efficacy construct, whereas the interest and outcome expectancy alphas were still too low (0.59 and 0.60, respectively). Again, further investigation is warranted as to the impact a conscientious or random responder has on the reliability of the constructs of this subscale.

The reliability coefficients for the engineering and technology subscale ranged from 0.49 for outcome expectations construct to 0.70 for the choice actions construct (Table 5). Only the choice actions construct achieved a satisfactory Cronbach's alpha value. However, the choice goals construct's Cronbach's alpha value is close to the satisfactory level. Like the mathematics and science subscales, if the reverse-worded items were removed, the self-efficacy and outcome expectations constructs achieve a

satisfactory level of Cronbach's alpha with the interest construct very close.

Like other surveys of this nature (Tyler-Wood et al. 2010; Unfried et al. 2015), the overall Cronbach's alpha coefficients were also calculated at the subscale level. The results yielded reliability coefficients of 0.89 for mathematics, 0.84 for science, and 0.87 for engineering and technology, indicating all items supported the subscale.

Content Validity

Content validity was used to examine the extent to which the items in the SIC-STEM Survey 2.0 are each aligned to one SCCT construct. As shown in Table 6, previously used working definitions from Roller et al. (2018) were developed for each of the SCCT constructs based on Lent (2013). These definitions were applied to each item that purported to measure a specific domain within the SCCT construct in question. Five researchers with specialties in mathematics education, science education, engineering/technology, or human development coded the SIC-STEM Survey 2.0 items to the five SCCT constructs after revising some of the initial SIC-STEM Survey items. The coding between the five researchers demonstrated 96% inter-rater reliability. Out of 45 items, four items demonstrated 60% inter-rater reliability, and three items demonstrated 40% inter-rater reliability. These were the only items not to score 100% inter-rater reliability. The seven items that demonstrated disagreement were discussed, and consensus was reached among all researchers that each item adhered to the definition before it was included in the final instrument for validation.

Table 4 Science subscale reliability coefficients

Construct	Cronbach's alpha	Cronbach's alpha if item deleted		
Interest	0.61*	S1 (0.48)	**S6** (0.59)	S11 (0.47)
Self-efficacy	0.56*	S2 (0.35)	S7 (0.30)	**S12** (0.74)
Outcome expectations	0.37*	S3 (0.18)	S8 (0.00)	**S13** (0.60)
Choice goals	0.82	S4 (0.75)	S9 (0.76)	S14 (0.73)
Choice actions	0.83	S5 (0.77)	S10 (0.77)	S15 (0.76)

S8 indicates science item 8. Bold denotes reverse-worded. * indicates a Cronbach's alpha below 0.7

Table 5 Engineering and technology subscale reliability coefficients

Construct	Cronbach's alpha	Cronbach's alpha if item deleted		
Interest	0.51*	**ET1** (0.68)	ET6 (0.21)	ET11 (0.25)
Self-efficacy	0.51*	**ET2** (0.71)	ET7 (0.19)	ET12 (0.28)
Outcome expectations	0.49*	ET3 (0.19)	ET8 (0.12)	**ET13** (0.75)
Choice goals	0.68*	ET4 (0.51)	ET9 (0.77)	ET14 (0.38)
Choice actions	0.70	ET5 (0.67)	ET10 (0.52)	ET15 (0.64)

ET2 indicates engineering and technology item 2. Bold denotes reverse-worded, * indicates a Cronbach's alpha below 0.7

Construct Validity

One of the most important features of scientific studies includes measuring and relating the variables, as well as revealing the causality (if any). However, some variables such as intelligence, interest, motivation, and achievement cannot be

Table 6 SCCT construct definitions

SCCT construct	Working definition	Item examples from SIC-STEM Survey 2.0
Self-efficacy	Items focus on one's perceived ability or a judgment of one's ability. For example, "I can do something" or "I am good/bad at something."	• I am not confident in math. • I can do well in science.
Outcome expectations	Items focus on the cause (decision or action) and effect (consequence of that decision). For example, "If I do this, then this will happen."	• Knowing how to design things will help me in the real world. • Doing problems helps me in math.
Interests	Items focus on whether students either like or do not like something. For example, "I like/do not like something."	• I like doing experiments. • I like to know how machines work.
Choice goals	Items focus on a goal or "wanting to do something" that was aligned to a STEM activity or career.	• I want to take math in college. • I want to choose a job in science.
Choice actions	Items need to 1) reference a career/job and 2) include a statement about taking an action. For example, "I work on problem solving skills in math club because I want a math job someday" implies that the decision to work on problem solving skills (an action) is viewed as being supportive toward getting a math job (future career goal).	• I use computers because I will need those skills in my job. • I ask for help when I need it because I will use math in my job.

Working definitions from "Title" by Roller et al. (2018)

directly measured. A model is proposed here to measure some of these variables (SCCT constructs). Confirmatory factor analysis was performed to evaluate the construct validity of the SCCT constructs. In this case, it is important to establish mathematical relationships that show how endogenous structures (predicted-endogenous) are linked with exogenous structures (predictive-exogenous; Yilmaz et al. 2006), through an analytical approach such as structural equation modeling (SEM; Hair et al. 1998), which consists of a combination of multivariate statistical techniques.

SEM is an effective way to build and test theoretical models that can explain possible causal relationships. SEM was used to investigate the 45 items aligned with the subscales (mathematics, science, and engineering and technology) and the constructs (interests, self-efficacy, outcome expectations, choice goals, and choice actions), which was the goal of the instrument's design. Furthermore, SEM uses a specified association structure to analyze correlations between observed measurements for the purposes of explaining model variability and uncovering relationships between latent variables. The software SmartPLS 3.0 was employed to explore relationships between unobserved constructs using partial least squares structural equation modeling (PLS-SEM).
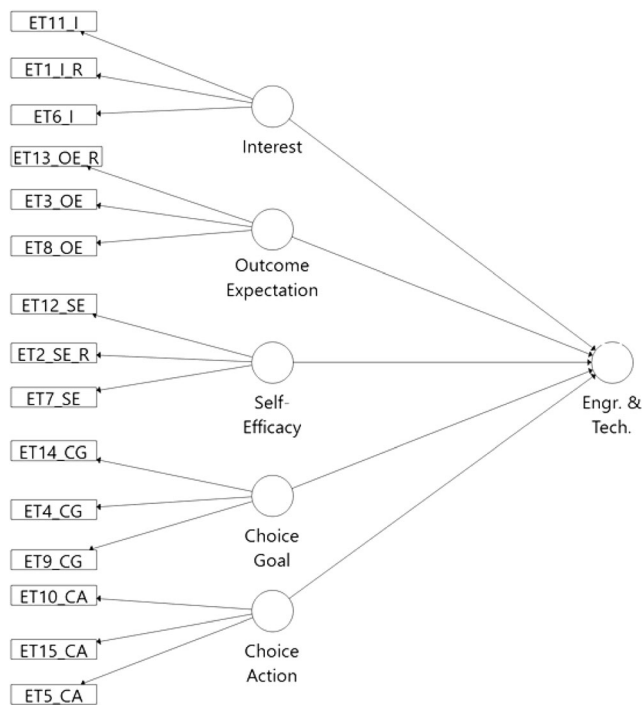
A model for each subscale (mathematics, science, and engineering and technology) was developed around the five SCCT constructs (interests, self-efficacy, outcome-expectations, choice goals, and choice actions). Figure 1 shows the SEM for the engineering and technology subscale. The SEM has three columns: from left to right the columns are the survey items, SCCT constructs, and subscale. Each survey item is further identified with abbreviations that represent the subscale, the item number, and the SCCT construct. Additionally, for a limited number of instances where the items were reverse-worded, an R is noted. Thus, ET13_OE_R represents engineering and technology item 13, measuring outcome expectations, and the item is reverse-worded. More importantly, the arrows in Fig. 1 suggest the direction of causal relationships between items and constructs. In this analysis, causality is only implied by correlation. This can be observed in the structural path diagram (Fig. 1) by the arrows pointing from latent constructs to

Fig. 1 Structural path diagram for engineering and technology subscale

observed items, which suggests the items are considered to be reflective of the constructs. The SEM for the science subscale and mathematics subscale is similar to Fig. 1.

Model Adequacy There are several different goodness-of-fit statistics researchers used to assess the adequacy of the structural equation models. The following are two of the measures suggested by Kline (2011) that could be calculated using SmartPLS 3.0.

1). The model chi-square (χ^2) is an assessment of the overall model fit. Chi-square is sensitive to sample size. The null hypothesis is that the model fits perfectly, so p values greater than 0.05 are desired.
2). Standardized root mean square residual (SRMR) is another measure of model fit. It is the square root of the difference between the residuals of the sample covariance matrix and the hypothesized model. SRMR is a metric often used when items vary in range (e.g., Likert Scale 1–5). SRMR values less than 0.08 are suggestive of an adequate fit.

We also used R^2, another goodness-of-fit metric, which measures the proportion of total variability explained by the model. Table 7 displays the results of the SEM for the analysis of the SIC-STEM survey for mathematics, science, and engineering and technology.

The p values for the goodness-of-fit chi-square test are all less than 0.05, which is below the established threshold for this metric. Due to the chi-square's sensitivity to sample size, it is recommended to use other model fit indexes, particularly for sample sizes over 400 where the chi-square statistic is almost always significant (Carvalho and Chima 2014). In the table above, the chi-square's p values are all less than 0.01, indicating that the null hypothesis of a "good" model fit would be rejected. This is due to the large sample size, and therefore, other fit measures will be used to evaluate the fit of the model.

The SRMR for the model indicates a good fit when the SRMR is less than 0.08, and an acceptable fit when the SRMR is less than 0.10 (Kline 2011). Each of the subscale models is on the threshold of being adequately fit with the mathematics subscale having a SRMR value of 0.09, the science subscale having a SRMR value of 0.12, and the engineering and technology subscale having a SRMR value of 0.11. For the mathematics model, a value of 0.09 indicates an acceptable fit, and the other two models are approaching an acceptable fit of 0.12 and 0.11.

The coefficient of determination (R^2) is the proportion of the total variability explained by the model. Hence, R^2 can be used to quantify the predictive accuracy of the model with respect to each latent subscale (Henseler et al. 2009). An R^2 value of at least 0.75 is considered to represent substantial model fit, while $R^2 \geq 0.50$ is considered to indicate a moderate fit, and $R^2 \geq 0.25$ is considered to indicate a weak fit (Henseler et al. 2009). There is moderate to substantial model accuracy across all three subscales with mathematics achieving an R^2 of 0.84, science an R^2 of 0.73, and engineering and technology having an R^2 of 0.82. While science is slightly below the recommended value, it is approaching a substantial fit, and thus, holistically, the mathematics and engineering and technology values demonstrate a substantial or good fit. Interpreting these values, for example, the five SCCT constructs explain 84% of the variability in the mathematics scores. Said another way, 84% of the variability in the mathematics scores can be attributed to the five SCCT constructs.

Table 7 SEMs goodness-of-fit measures

Measure	Mathematics	Science	Engineering and technology	Cutoff for good fit
χ^2	986.25 ($p < 0.01$)	1327.92 ($p < 0.01$)	1203.72 ($p < 0.01$)	$p > 0.05$
SRMR	0.09	0.12	0.11	SRMR < 0.08
R^2	0.84	0.73	0.82	$R^2 > 0.75$

In summary, the goodness-of-fit indices demonstrate that the model is acceptable. The empirical evaluation of the model included evaluating the fit indices χ^2 and SRMR, and the R^2 value implies acceptable fit.

Model Analysis Since the model demonstrated acceptable fit, the model was analyzed to investigate the relationships among the constructs and subscales. The effects of the constructs (interests, self-efficacy, outcome-expectations, choice goals, and choice actions) on the subscales (mathematics, science, and engineering and technology) were investigated using f^2 to determine what impact each construct had on each subscale. The path coefficients of the survey items on the constructs were also considered to determine the strength of association between the different structures in the model.

The f^2 (the sampling distribution to measure each construct's effect on each subscale) statistic measures the degree of the impact one latent variable has on another, which means it measures the strength of each predictor variable in explaining the endogenous variables (science, mathematics, engineering and technology). This is determined by measuring how much of the explained variability is unique due to a particular variable. According to Cohen (1988), an f^2 value of 0.350 or greater is representative of a strong effect, at least 0.150 a moderate effect, and at least 0.020 a weak effect. Table 8 displays the strengths of each construct with respect to each subscale. The results indicate that there are no strong effects, with most of the effects being between moderate and weak (Table 8). This implies that all constructs influence the subscales to some degree; however, some of these influences are weak, and some are moderate according to Cohen. Taking the moderate f^2 statistics into consideration, students' intent to pursue a STEM career would be most influenced by students' interest, outcome expectations, and choice goals regarding mathematics; students' self-efficacy in science; and students' interests and outcome expectations regarding engineering and technology. For example, that would mean increasing students' choice goals regarding mathematics (wanting to do something that was aligned to a STEM activity or career)

Table 8 f^2 statistics for each construct and schedule

| Construct	Mathematics	Science	Engineering and technology
Interest	0.171*	0.079	0.191*
Self-efficacy	0.067	0.137*	0.074
Outcome expectations	0.214*	0.148	0.155*
Choice goals	0.155*	0.057	0.109
Choice actions	0.070	0.081	0.125

*indicates moderate effect
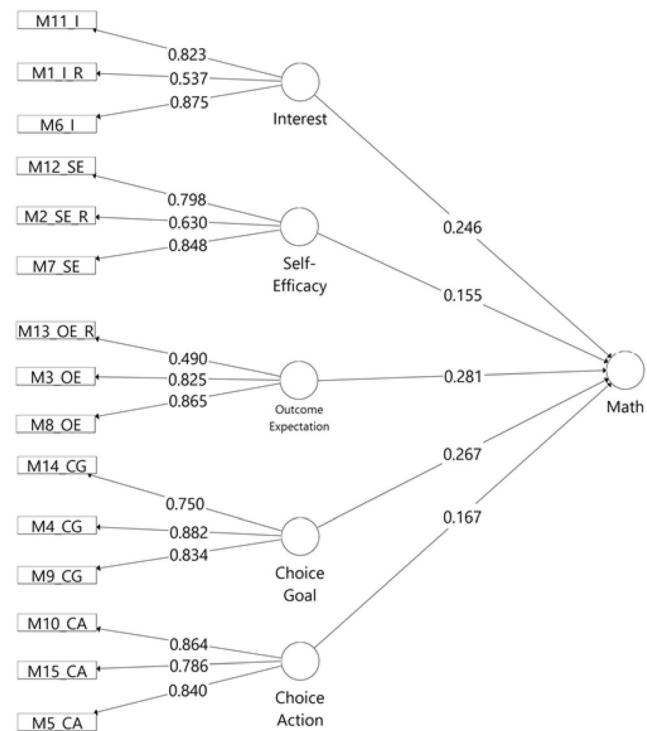
Fig. 2 Mathematics subscale structural equation model

would have a moderately positive influence on students' pursuit of a STEM career.

The path coefficients describe the associations between the different structures in the model. These path coefficients (or connection strengths) are the parameters of the model and represent the estimates of effective connectivity of the path models shown in Figs. 2, 3, and 4. For example, 0.823 (see Fig. 2) is the path coefficient of the first mathematics interest
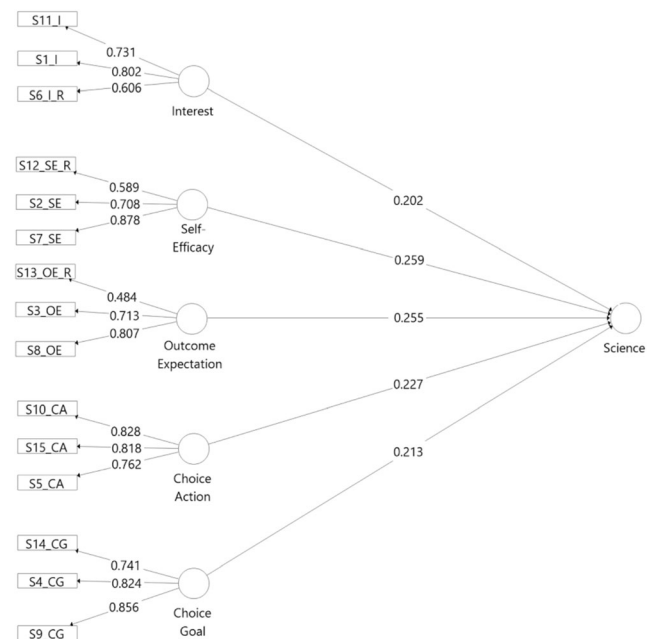
Fig. 3 Science subscale structural equation model

survey item 1. This implies if the mean of the mathematics interest survey item 1 increases by one standard deviation, then the mean of the interest construct would increase by 0.823; one standard deviation from its mean. In general, the path coefficients for the three models measure the increase in standard deviation on the latent variable if the observed variable is increased. Figures 2, 3, and 4 depict the path coefficients that were computed via the PLS-SEM algorithm for each subscale.

Lastly, the t values for each construct to each subscale were used to test the hypothesis that each construct had a significant effect on each subscale. All t values had a p value less than 0.001. This implies that the coefficients in the model were statistically significant at a level less than 0.001. This is interpreted to mean that each model construct had a statistically significant effect on each subscale. For example, all five constructs of interest, self-efficacy, outcome expectations, choice goals, and choice actions are useful in predicting the subscales of science, mathematics, and engineering and technology. This confirms construct validity for the SIC-STEM Survey 2.0.

Discussion

The SIC-STEM Survey 2.0 is an instrument that was intentionally designed to collect information about the factors that influence student STEM career interest and choice in regard to the five constructs (interests, self-efficacy, outcome expectations, choice goals, and choice actions) that are
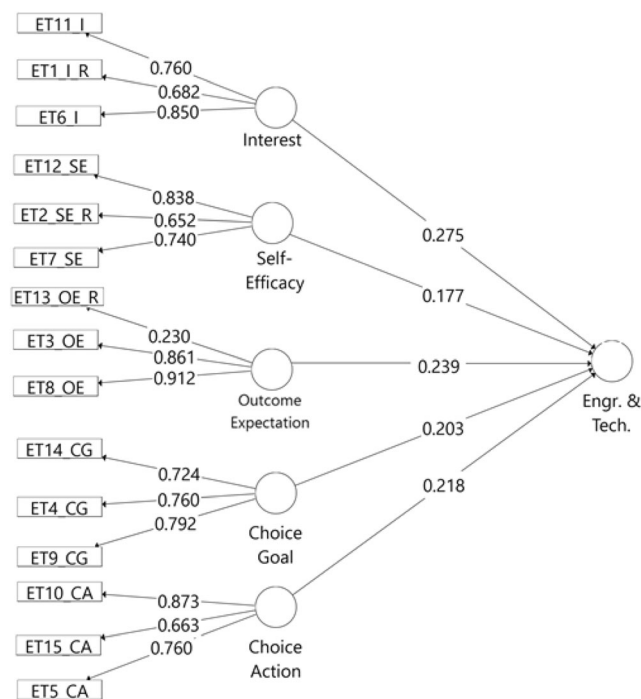

Fig. 4 Engineering and technology subscale structural equation model

utilized in the SCCT interest and choice models (Roller et al. 2018). While other instruments have aligned with the SCCT framework (S-STEM; Faber et al. 2013; Unfried et al. 2015; STEM-CIS; Kier et al. 2013), the SIC-STEM Survey 2.0 is the first survey instrument to collect data about choice goals and choice actions. This in turn provides information beyond students' interests in STEM, specifically how students leverage their decisions and choices toward a STEM career path. Data of this nature have the potential to be useful in understanding and addressing the national shortage of STEM graduates.

Instrument Reliability and Validation

This study supports the research questions that the revised SIC-STEM Survey 2.0 has the potential to produce reliable and valid data among 4th–12th grade students. Cronbach's coefficient alphas were calculated for each subscale (mathematics, science, and engineering and technology) and construct (interests, self-efficacy, outcome expectations, choice goals, and choice actions), which demonstrated levels of acceptable reliability coefficients in some of the constructs. Reverse-worded items also tended to weaken the construct reliability, with each subscale containing at least two items that, if deleted, would raise the Cronbach's alpha for that construct. A reverse-worded item appeared in the interest, self-efficacy, and outcome expectations constructs for each subscale. Further analysis is warranted as to the impact the reverse-worded items have on the detection of a conscientious or random responder. While the instrument could be improved at the item level, all three subscale composite reliability coefficients are at 0.84 or above, which implies that the data collected with the SIC-STEM Survey 2.0 demonstrates subscale reliability and only some construct reliability.

The instrument was analyzed further to determine if indeed it is an instrument that produced valid and reliable data with the targeted population. Content validity demonstrated a 96% inter-rater reliability among five researchers, with consensus being reached for 100% of the items, indicating the items were aligned to the construct definitions. Construct validity of items were investigated by utilizing PLS-SEM. The structural equation model analysis for each subscale in the SIC-STEM Survey 2.0 supported items was reflective of the SCCT constructs. Additionally, the goodness-of-fit indices (χ^2, SRMR, and the R^2 values) demonstrated that the model is acceptable. Further model analyses provided path coefficients, which suggest each construct item had a weak to moderate influence on the subscale. Particularly interesting was that different SCCT construct path coefficients appeared as moderately influential in each subscale, versus one or two SCCT constructs being stronger across all three subscales. Specifically, the path coefficient for choice actions in mathematics demonstrated a moderate influence on whether a student would pursue a

STEM career, which is important to highlight because previous instruments had not collected information regarding choice actions. Moreover, *t* values indicated that each construct had a statistically significant effect on the subscale, thus concluding that the five SCCT constructs are predictive of the subscale in the SIC-STEM 2.0 Survey.

Limitations

The large variability in the sampled students' ages (4th–12th grade) demonstrates that the SIC-STEM Survey 2.0 can be used to collect data across broad populations. As more data is gathered in future studies, it may be interesting to analyze the results in grade bands (4th–5th, 6th–8th, and 9th–12th) to highlight elementary, middle, and high school trends. The reverse-worded items were originally included in the instrument to identify whether the students were paying attention when completing the survey. While intended as a delimitation, Cronbach's alpha values indicate that the reverse-worded items were often detrimental to the instrument's reliability. Future studies may choose to include these items rewritten in a positive way to determine if they still load on their appropriate categories. Lastly, a limitation of the instrument is that it was designed with only three items per construct. The rationale for this design was to create a survey that fit the attention span of students and could be efficiently implemented in schools without consuming too much instructional time. Thus, while a practical recommendation to increase instrument reliability further would be to increase the number of items beyond three, SEMs with many constructs and items tend to not fit as well as SEMs with few constructs and items (MacCallum et al. 1996).

A limitation in the implementation of the SIC-STEM Survey 2.0 instrument was the researchers did not observe students taking the survey but instead had teachers administer the instrument. While directions were included with the instrument on how to administer the survey, teachers naturally try to support students and may have read items or even helped students understand items by providing them with definitions or examples. These definitions may have included inaccuracies or specific examples, which may have influenced students' opinions. Readability was also explored during survey development to promote accessibility; however, readability does not imply comprehension. Furthermore, the interpretation of the items by students in 4th to 12th grade was not investigated during this study. Thus, it could be that students were interpreting these items similarly or differently. This limitation could be explored in the future by completing student interviews in various grade levels to learn more about how students are comprehending and interpreting each item.

Areas for Future Investigation

Moreover, the reliability analysis of the SIC-STEM Survey 2.0 has highlighted two areas that need to be investigated further: (1) the impact conscientious or random responders have on the survey's psychometric properties and (2) the level of Cronbach's alpha that is deemed satisfactory for the results of this survey to be interpreted in a significant way. Because this survey was designed with a reverse-worded item for the interest, self-efficacy, and outcome expectations constructed for the three subscales, it provides an opportunity for a researcher using this survey to determine if the survey responder was behaving as a conscientious or a random responder. Several techniques (Curran 2016; Marjanovic et al. 2015) have been developed recently that need to be explored to determine their usefulness in detecting a conscientious or random responder and the impact those detections have on the reliability of this instrument. Data from the SIC-STEM Survey 2.0 may afford this sort of exploration and analysis.

The entirety of this survey's reliability analysis is based on a satisfactory level of Cronbach's alpha of 0.7 from Nunnally (1978). According to Cho and Kim (2015), Nunnally's work has become a standard used by many in their reliability assessment for an instrument. However, several researchers (Cho and Kim 2015; Schmitt 1996; Taber 2018) have recently begun to question the 0.7 level, since it was derived from Nunnally's personal intuition and not from empirical evidence (Churchill and Peter 1984; Peterson 1994). Therefore, as additional data is gathered and analyzed with this instrument, it is recommended that the acceptable level of alpha continues to be explored.

During validation testing, specific SCCT constructs within each subscale were illuminated as having more of an influence on students pursuing STEM careers. Specifically, students' interest, outcome expectations, and choice goals were highlighted in mathematics, students' self-efficacy in science, and students' interests and outcome expectations in engineering and technology. Thus, this data suggests it might be more important for science teachers to focus their efforts on supporting students' belief that they can do science (self-efficacy) than it would be to make the content more interesting, given that the self-efficacy construct had a stronger path coefficient. Moreover, mathematics had the most SCCT constructs in a subscale to be identified as moderately influential. Future research investigating the impact of mathematics, and specifically the SCCT constructs regarding mathematics, has on students' impact to pursue STEM careers could be valuable work. Exploring ways to strengthen interest, outcome expectations, and choice goals in mathematics may provide a productive space for widening the STEM career pipeline. Additionally, more generalizable studies about how to best impact a SCCT construct within any subject area could also

be beneficial for addressing the STEM workforce shortage earlier, perhaps with STEM experiences in K-12 education.

Appropriate Uses

Based on this study, the SIC-STEM Survey 2.0 has demonstrated that the instrument can collect valid and reliable data from a broad range of students, specifically enrolled in 4th to 12th grade. Since the survey relies heavily on a student's ability to read and comprehend the text, this instrument is not recommended for implementation with early elementary students (K-3rd grade) but could perhaps be effective at gathering data from undergraduate students. This extension into post-secondary education could provide insights into the differences and similarities between engineering and non-engineering students' STEM interest and choice. Future studies investigating interest and choice in post-secondary education might also extend work from childhood through college.

The instrument has been implemented within the educational environment because it is a space where students feel safe and a teacher can easily assist if a student has difficulty understanding the survey or becomes overwhelmed completing the survey. Situations of this nature were not reported during this study; however, it would be important to have a responsible adult who is able to respond to a student (e.g., reading a word, stopping the survey) in these types of situations. Furthermore, because the survey length is already substantial in order to collect data about each construct and subscale, it is recommended that this instrument only be combined with a limited number of additional items (e.g., demographics, career interest), similar to this study. If researchers want to gather data beyond the SCCT constructs for STEM subscales that require more lengthy instruments, then it would be recommended to employ these instruments in a separate sitting to help break up the time.

Future Uses of the Instrument

The SIC-STEM Survey 2.0 demonstrated satisfactory levels of reliability and validity suggesting it can be utilized for larger data collection with students. Specifically, this instrument could provide individual schools with data that might aid administrators in understanding the state of their students' interests and choices toward STEM. Having this knowledge would provide school system personnel with information to guide ongoing curriculum development, as well as to encourage partnerships with STEM employers and stakeholders. Furthermore, by using data from this instrument to modify or create new opportunities for students that are intentionally designed with a goal of strengthening more influential SCCT constructs in STEM subjects, more students might pursue STEM fields and careers.

Compliance with Ethical Standards

References

Adelson, J. L., & McCoach, D. B. (2011). Development and psychometric properties of the math and me survey: measuring third through sixth graders' attitudes toward mathematics. *Measurement and Evaluation in Counseling and Development, 44*(4), 225–247.

American College Testing (ACT). (2017). STEM education in the U.S.: Where we are and what we can do. Retrieved from: www.act.org/content/dam/act/unsecured/documents/STEM/2017/STEM-Education-in-the-US-2017.pdf

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

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

Burton, L. J., & Mazerolle, S. M. (2011). Survey instrument validity part 1: principles of survey instrument development and validation in athletic training education research. *Athletic Training Education Journal, 6*(1), 27–35.

Capobianco, B. (2015). Engineering identity development scale (EIDS). *Purdue University Research Repository*. https://doi.org/10.4231/R7WH2MX6.

Carvalho, J., & Chima, F. O. (2014). Applications of structural equation modeling in social sciences research. *American International Journal of Contemporary Research, 4*(1), 6–11.

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: well known but poorly understood. *Organizational Research Methods, 18*(2), 207–230.

Churchill Jr., G. A., & Peter, J. P. (1984). Research design effects on the reliability of rating scales: a meta-analysis. *Journal of Marketing Research, 21*(4), 360–375.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology, 66*, 4–19.

Erkut, S., & Marx, F. (2005). 4 schools for WIE (evaluation report). Wellesley: Wellesley College, Center for Research on Women.

Faber, M., Unfried, A., Wiebe, E. N., Corn, J., Townsend, L. W., & Collins, T. L. (2013). Student attitudes toward STEM: The development of upper elementary school and middle/high school student surveys. In *Proceedings of the 2013 American Society for Engineering Education annual conference & exposition*. Washington, DC: ASEE.

Germann, P. J. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal*

of Research in Science Teaching, 25(8), 689–703. https://doi.org/10.1002/tea.3660250807.

Guzey, S., Moore, T., & Harwell, M. (2014). Development of an instrument to measure students' attitudes toward STEM. *School Science and Mathematics, 114*(6), 271–279.

Hair, J. H., Tatham, R. L., & Anderson, R. E. (1998). *Multivariate data analysis* (5th ed.). New York: Prentice Hall International.

Henseler, J., Ringle, C. M., & Sinkovics, R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing: Vol. 20. New Challenges to International Marketing* (pp. 277-319). https://doi.org/10.1108/S1474-7979(2009)0000020014.

Kennedy, J., Quinn, F., & Taylor, N. (2016). The school science attitude survey: a new instrument for measuring attitudes towards school science. *International Journal of Research & Method in Education, 39*(4), 422–445. https://doi.org/10.1080/1743727X.2016.1160046.

Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2013). The development of the STEM career interest survey (STEM-CIS). *Research in Science Education, 44*(3), 461–481.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

Lent, R. W. (2005). A social cognitive view of career development and counseling. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: putting theory and research to work* (pp. 101–127). Hoboken, NJ: John Wiley & Sons Inc.

Lent, R. (2013). Social cognitive career theory. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: Putting theory and research to work* (pp. 115–146). Hoboken: Wiley.

Lent, R. W., & Brown, S. D. (2006). Integrating person and situation perspectives on work satisfaction: a social-cognitive view. *Journal of Vocational Behavior, 69*(2), 236–247.

Lent, R. W., & Brown, S. D. (2008). Social cognitive career theory and subjective well-being in the context of work. *Journal of Career Assessment, 16*(1), 6–21.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social-cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior, 45*(1), 79–122.

Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: a social cognitive analysis. *Journal of Counseling Psychology, 47*(1), 36–49.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149.

Marginson, S., Roberts, K. & Tytler, R. (2013). *STEM: Country comparisons: International comparisons of science, technology, engineering, and mathematics (STEM) education. Final Report.* Australian Council of Learned Academies, Melbourne, VIC.

Marjanovic, Z., Holden, R., Struthers, W., Cribbie, R., & Greenglass, E. (2015). The inter-item standard deviation (ISD): an index that discriminates between conscientious and random responders. *Personality and Individual Differences, 84*, 79–83.

National Research Council. (2011). *Successful STEM education: A Workshop summary.* A. Beatty, Rapporteur. Committee on Highly Successful Schools or Programs for K-12 STEM Education, Board on Science Education and Board on Testing and Assessment. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Noonan, R. Office of the Chief Economist, Economics and Statistics Administration, U.S. Department of Commerce. (2017). *STEM Jobs: 2017 update (ESA Issue Brief # 02–17).* Retrieved from https://files.eric.ed.gov/fulltext/ED594354.pdf

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21*(2), 381–391.

Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review, 29*(6), 901–910.

Roller, S. A., Lampley, S. A., Dillihunt, M. L., Benfield, M. P., & Turner, M. W. (2018). Student attitudes toward STEM: a revised instrument of social cognitivecareer theory constructs (Fundamental). In *Proceedings of the 2018 American society for engineering education annual conference & exposition.* Salt Lake City, UT: ASEE.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350–353.

Scholtes, V. A., Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury, 42*(3), 236–240.

Shapiro, C. A., & Sax, L. J. (2011). Major selection and persistence for women in STEM. *New Directions for Institutional Research, 2011*(152), 5–18.

Taber, K. S. (2018). The use of cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296.

Tapia, M., & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly, 8*(2), 16–21.

The President's Council of Advisors on Science and Technology (PCAST). (2012). Report to the president—*Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Washington, DC: U.S,. Government Printing Office. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education, 18*(2), 345–368.

Unfried, A., Faber, M., Stanhope, D. S., & Wiebe, E. (2015). The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *Journal of Psychoeducational Assessment, 33*(7), 622–639.

Washington, A.N., Grays, S. & Dasmohapatra, S. (2016). The computer science attitude and identify survey (CSAIS): A novel tool for measuring the impact of ethnic identity in underrepresented computer science students. Presented at the *American Society for Engineering Education*, New Orleans, LA.

Wilson, R.A., Sofroniou, N., Beaven, R., May-Gillings, M., Perkins, S., Lee, M. ... Pirie, J. (2016). *Working Futures 2014-2024: UKCES evidence report no 100.* Retrieved from https://www.gov.uk/government/publications/uk-labour-market-projections-2014-to-2024

Yilmaz, V., Celik, H. E., & Ekiz, E. H. (2006). Investigation of the factors that affect the Authority's commitment with structural equation modeling; a sample of bank of primary and government. *Journal of Social Sciences, 2*, 171–184.