Students Attitudes Surrounding STEM: A Social Cognitive Career Theory Instrument for High School

EmilyKate McDonough[1,2], Kayle S. Sawyer[3,4,5,6], Jessica Wilks[7], and Berri Jacque[1]

[1] Tufts University School of Medicine, Boston, MA, USA

[2] Harvard Medical School, Boston, MA, USA

[3] VA Boston Healthcare System, Boston, MA, USA

[4] Boston University School of Medicine, Boston, MA, USA

[5] Massachusetts General Hospital, Boston, MA, USA

[6] Sawyer Scientific, LLC, Boston, MA, USA

[7] Resilient Games Studio, LLC, Chicago, IL, USA

Corresponding Author: EmilyKate McDonough, Ph.D.

Tufts University School of Medicine, Sterns 704

136 Harrison Ave

Boston, MA 02111, USA

Phone: +1 617-636-6913

Email: emilykate.mcdonough@tufts.edu

Keywords: STEM, SCCT, instrument, validation, reliability

Running Title: SASS Instrument Validation

Abstract

To meet the demand of the growing science and health sectors in the United States, there is a critical necessity to engage more people in science, technology, engineering, and math (STEM). To broaden participation in STEM, we must understand the factors that shape perspectives and beliefs around career selection. Good measurement of these factors is crucial to quantify how effectively educational interventions impact student attitudes towards STEM. Adolescents are particularly suited for quantifying intervention efficacy because students build their identities during these formative years and make important career choices. To better quantify intervention efficacy at the high school level, we developed an instrument entitled Student Attitudes Surrounding STEM (SASS), which builds upon the social cognitive career theory (SCCT) framework for understanding career selection. Questionnaire responses were collected from 932 high school students, and split into samples of 400 for exploratory factor analysis and 532 for confirmatory factor analysis. The questions clustered into six factors: self-efficacy experience, self-efficacy academic, outcome expectations, interests, negative perceptions of scientists, and career awareness. The SASS exhibited adequate construct validity as determined by fit indices and theoretical considerations. Furthermore, the instrument demonstrated criterion validity, internal consistency, and test-retest reliability. This tool represents a novel integration of three latent variables into SCCT, negative perceptions of scientists, career awareness, and an experience factor for self-efficacy.

Introduction

Educating students in science, technology, engineering, and math (STEM) is essential for their development into informed citizens and active members of the 21st century workforce. Additionally, the skills used in STEM help people lead fulfilling lives, make sensible personal decisions, and support productive civic and community engagement. However, there is growing concern of a 'leaky STEM pipeline,' referring to the high rate of attrition from STEM majors and graduate degrees. Over the past couple of decades in the United States, 40%-50% of undergraduate students who enrolled in STEM majors left the field before their scheduled graduation, or shortly thereafter (Chen, 2013; PCAST, 2012). Particularly high rates of attrition have been seen in the populations already underrepresented in STEM, such as women, people of color, and first generation college students (Anderson & Kim, 2006; Gayles & Ampaw, 2014; Hill et al., 2010; Shaw & Barbuti, 2010). This loss of talent within the STEM workforce leads to the reduction of diverse perspectives and ideas, and limited representation of community stakeholders that are essential for progress in science and medicine.

Just as the paths people take to obtain a STEM career are varied, so too are their reasons for exiting the STEM pipeline. Numerous contextual factors, such as access to role models, peer support, institutional environment, and discrimination contribute to a person's decision to leave STEM (Chang et al., 2011; Fouad et al., 2010; Gayles & Ampaw, 2014; Griffith, 2010; Hill et al., 2010; Wang, 2013). For example, schools that focus their teaching and research efforts on undergraduates, rather than graduate students, retain more students within the STEM majors (Griffith, 2010). Additionally, various individual factors impact which students are likely to leak out of the

STEM pipeline. For example, undergraduate students majoring in STEM fields tend to switch to a non-STEM major if they have lower confidence in their STEM abilities, and this is especially true for women and first generation college students (Shaw & Barbuti, 2010; Wang, 2013). Another factor linked to withdrawal from STEM comes from a waning interest in the field, due to a lack of knowledge surrounding potential STEM careers (Cohen et al., 2013). If students don't know what careers are available, or what a STEM career actually entails, they may be less likely to pursue one.

While the leaky pipeline term has been criticized for its suggestion that there is a single trajectory to a STEM career (Cannady et al., 2014), the imagery evoked by the term has increased awareness of the attrition problem and prompted the development of numerous STEM education interventions worldwide (UMass Donahue Institute, 2011; van den Hurk et al., 2019). Interventions that effectively retain students in STEM build knowledge and influence student trajectories by fostering healthy self-perceptions and beliefs. To quantify the success of these educational interventions, instruments measure characteristics such as domain knowledge, extracurricular knowledge, aspects of self-perception, and aspirations. Data from these instruments can identify which student groups benefit from the intervention, allowing an educational program to be implemented for those with the greatest need and for those who would receive the greatest benefit. Finally, ideal metrics reveal the efficacy of the intervention in "real time", allowing for iterative curricular improvements.

In summary, to broaden participation in STEM, careful measurement of educational interventions' impact on student perspectives is essential. Adolescence is a critical period in which to assess interventions, because students form beliefs about themselves and make enduring choices regarding career selection. In this manuscript

we describe the development and validation of a new questionnaire, which we call Student Attitudes Surrounding STEM (SASS).

Background and Theoretical Framework

Definition of STEM

There are many definitions of STEM careers, each including a different set of jobs. For example, some include healthcare work within the realm of STEM, while others do not (Koonce et al., 2011). In this manuscript, we define STEM careers broadly as any career that requires knowledge from a STEM domain. Thus, we include all health-related jobs, as well as careers that require data science (e.g., psychology research), as well as information technology and computer science. We only include jobs that directly involve the fields of science, technology, engineering, or math, which excludes jobs such as counseling. We selected a broad definition of STEM so that our research would encompass a full range of current and emerging STEM roles that may be applicable to high school students' diverse interests.

Social Cognitive Career Theory

In 1994, Robert W. Lent, Steven D. Brown, and Gail Hackett proposed the social cognitive career theory (SCCT) as a unified framework through which to understand career choices. The theory is based on the work of Albert Bandura's social cognitive theory, which examines the relationship between personal factors (e.g., beliefs, attributes), environmental factors (e.g., feedback, culture), and behavior (Bandura, 1986). At its core, SCCT measures the relationships between self-efficacy ("I am good at X"), outcome expectations ("If I do X, this will happen"), interests ("I'm interested in X"), and goals ("I plan to do X"), as described by Lent and colleagues (1994, 2000). For

example, a STEM-specific model suggests that someone who considers themselves good at math, and who expects positive outcomes from a math-related career, would be more interested in math, set goals to study math, and persist within STEM. That is, higher confidence and belief in favorable outcomes will help students set goals and persist within STEM. Thus, self-efficacy, outcome expectations, and interests, influence career-related goal-setting and behavior.

While the four cognitive factors (self-efficacy, outcome expectations, interests, and goals) represent the center of SCCT, they do not exist in isolation. People respond to their environments and incorporate social feedback as they develop beliefs about themselves. Therefore, researchers have examined how additional constructs influence these core factors, and ultimately, career choice. Lent and colleagues (1994, 2000) proposed that personal characteristics (e.g., gender, race/ethnicity, emotional affect, socioeconomic status [SES]), background factors (e.g., family expectations, role model exposure), and learning experiences influence these cognitive factors and mediate their impact on career choice. Additionally, career choice outcomes are influenced by external factors such as the supports (e.g., social network) and barriers (e.g., discrimination) one might expect to receive as a consequence of choosing a STEM career (Lent et al., 1994, 2000). While no single environmental element determines outcomes by itself, each can have a powerful impact. For example, high SES, which allows for greater access to learning experiences and career-support resources, serves as one form of support that influences career choices and persistence. However, people with financial resources do not always succeed in attaining their career goals, whereas people with few resources have succeeded none-the-less. People respond to their

environment differently for a wide variety of reasons, as captured by the various environmental inputs of the SCCT framework.

SCCT has been used to examine STEM career choice in a wide range of populations. The framework has been used for research examining outcomes for low income, prospective first generation college students (Garriott et al., 2013), as well as for examining effects of race, ethnicity and gender (Lent et al., 2005; Navarro et al., 2007; Turner et al., 2019). These individual characteristics are commonly referred to as *person inputs*. SCCT has also been used to examine populations with varying SES, although findings have been inconsistent. One study within a population of Mexican-American adolescents found that SES influences outcome expectations, but not self-efficacy (Turner et al., 2019). Other research with a mixed student sample from the mid-western United States shows SES affects self-efficacy, but not outcome expectations (Navarro et al., 2007). While mostly explored at the undergraduate and graduate level, the SCCT model has been validated for various age groups, including high schoolers, college students, and people early in their careers (Fouad & Santana, 2017; Lent et al., 2018). Indeed, the SCCT framework is a powerful tool for understanding career choice among different populations. However, the relationships between the SCCT constructs have varied when examining different populations, especially with respect to self-efficacy and outcome expectations (Fouad & Santana, 2017). These studies illustrate the framework's utility, while also highlighting the need to validate the SCCT model within specific contexts.

People begin to consider their future professions starting from a young age, often before beginning their formal education. Throughout childhood and adolescence changes in self-efficacy, outcome expectations, and interests can impact career choice.

Adolescence is a crucial time in people's lives with respect to their career, because it is when people make impactful decisions, such as whether to attend college. Measuring the SCCT constructs in high school students allows us to understand what motivates potential STEM professionals early in their career trajectory. SCCT research within the high school population has focused on career interest as an endpoint measure, rather than career goals or actions (Lent et al., 2018). This may reflect how interests are a more pertinent measure for adolescents than goals, which may not be well-formed in this population. Science self-efficacy has been correlated with science career interest for students in health career-focused extracurricular programs, although this observation may be specific to the selected population of adolescents who show high interest and achievement in STEM (Peterman et al., 2018). However, even as early as middle school, positive experiences in science and math are associated with higher STEM self-efficacy and career interest (Fouad & Santana, 2017). These studies show that researchers can use SCCT to robustly measure a STEM intervention's impact on future career choices in adolescent populations.

The literature examining the SCCT latent variable relationships during adolescence has mainly focused on single subjects such as math (Garriott et al., 2013; Lim & Chapman, 2013; Lopez et al., 1997; Lopez & Lent, 1992; Schukajlow et al., 2012) or single disciplines, such as science (Stake, 2006; Syed et al., 2012). To our knowledge, only three instruments have been developed to investigate the SCCT constructs in the context of all-inclusive STEM with high school students: the STEM Career Interest Survey (STEM-CIS; Kier et al., 2014), the Student Attitudes Toward STEM (S-STEM; Unfried et al., 2015), and the Student Interest and Choice in STEM (SIC-STEM; Roller et al., 2020). STEM-CIS was used to assess Taiwanese high school

STEM career interest in the context of SCCT, but it was validated using the discipline-specific scales (science, technology, engineering, and math), rather than the SCCT constructs (self-efficacy, outcome expectations, interests, goals, supports, and person inputs; Kier et al., 2014). Likewise, S-STEM, and its derivative SIC-STEM, was developed with individual scales for 21st century skills, science, math, and a combined scale for science and technology (Roller et al., 2020; Unfried et al., 2015). Recently, STEM-CIS was adapted for high school students in China and validated using the SCCT scales, demonstrating that the instrument could be used for evaluation of the SCCT constructs (Mau et al., 2019). While the 44-item STEM-CIS instrument could be used for measuring the key SCCT constructs within a high school population, it does not include items that quantify perceptions of scientists, nor career awareness, which we were interested in measuring.

Perceptions of Scientists

In 1957, Margaret Mead and Rhoda Métraux asked approximately 35,000 high school students across the United States to write essays about their views on scientists. Students responding to the writing prompt often described a scientist's work as admirable and valuable to society. However, this positive view was paired with negative depictions of a scientist when students were asked to describe themselves or a romantic partner as a scientist. For example, one student described a scientist as having "no other interests" and as someone who "neglects his family" (Mead & Métraux, 1957). Research built upon this work found that several of these false stereotypes have persisted for decades (Chambers, 1983; Dikmenli, 2010; Finson, 2002; Pion & Lipsey, 1981; Schinske et al., 2015; Scholes & Stahl, 2020; Wyer et al., 2010). Furthermore, students can be ignorant about what scientists actually do, believing that all scientists

conduct dangerous work and shout “Eureka!” upon successful completion of an experiment (Chambers, 1983; Christidou, 2011; Mead & Métraux, 1957; Scholes & Stahl, 2020). These stereotypes of scientists and science have been observed in students across the globe and in various settings (Christidou, 2011). In summary, research spanning 60 years shows that secondary students hold a confused and unrealistic picture of what it means to be a scientist.

The persistence of negative perceptions of scientists is especially concerning as it dissuades people from pursuing STEM, which in turn contributes to the leaky pipeline. Indeed, stronger aspirations of undergraduate students to pursue a STEM career are associated with more positive perceptions of scientists (Wyer, 2003). Furthermore, science self-efficacy and interests (major influencers in career choice) are interconnected with student-held images of scientists (Christidou, 2011). Negative stereotypes of scientists are seen more prevalently in Asian undergraduate students than in their peers (Schinske et al., 2015), demonstrating a connection between person inputs (in this case, race) and perceptions of scientists. Interventions that paint a realistic picture of what it means to be a scientist show promise in addressing the racial and gender disparities within STEM.

Developing good measures of students' stereotypes of scientists is crucial to assess the impact of educational interventions aimed at changing the perception of scientists. Several methods have been used to accomplish this goal, ranging from prompted essays (e.g., Mead & Métraux, 1957) and drawings (e.g., Chambers, 1983) to Likert scale rated questionnaires (Krajkovich & Smith, 1982; Wyer et al., 2010). The questionnaires developed by Krajkovich and Smith (1982) and Wyer and colleagues (2010) included items inspired by imagery and beliefs detailed in the initial work of Mead

and Métraux (1957). Both scales included positive (e.g., intelligent and careful) and negative (e.g., unhappy family and limited social life) stereotypes of scientists. While the survey developed by Krajkovich and Smith (1982) was initially specified with a single factor, later work highlighted the independence of negative perceptions (Marshall et al., 2010). These lengthy instruments help to uncover the breadth of stereotypes students have about scientists, which can be useful in designing interventions that target specific beliefs. To date, research has generally examined perceptions of scientists independently from SCCT constructs (Christidou, 2011). Integrating the two measures would help researchers understand students' choices to persist within, or exit from, STEM and may reveal connections between perceptions of scientists and self-efficacy, outcome expectations, interests, or goals.

Career Awareness

It is hard to develop an interest in a career if one is unfamiliar with it. While most students are familiar with STEM careers such as doctor or computer programmer, fewer are aware of more specialized jobs such as animal care technician or clinical trial recruiter. Within the high school population, higher STEM career awareness, along with higher STEM self-efficacy, is correlated with stronger desires to pursue a STEM career, demonstrating an interplay between SCCT and career awareness (Blotnicky et al., 2018; Zhang & Barnett, 2015). Furthermore, a study of students from a primarily white, middle class high school showed that adolescents are unaware of the skills and knowledge required for entry into their chosen fields of interest (Johnson, 2000). Assuming that career awareness is associated with privilege, then under-resourced and minority groups would have even less awareness than the white, middle class students.

This research highlights the importance of incorporating career-specific knowledge into curricular interventions aimed at increasing interest in STEM careers.

Career knowledge can be quantitatively assessed by directly measuring one's knowledge of required skills, degrees, or duties for specific careers. For example, the Cognitive Vocational Maturity Test (CVMT) is a highly comprehensive, 120-item, six-construct instrument that assesses career knowledge in middle and high school students. While the CVMT provides a wealth of information about an individual's career knowledge, the full instrument takes two 45-minute class periods to complete, and is therefore impractical for most settings (Westbrook et al., 1996). Alternatively, some instruments measure specific knowledge about careers, such as what degree or classes are required to enter particular fields (e.g., Blotnicky et al., 2018). Another approach to understanding career awareness and career-related knowledge is to rely on self-reported learning. For example, Salonen and colleagues (2018) used three questions to assess perceived learning following a career awareness intervention: *"I gained knowledge about careers that are new to me," "This unit helped me understand the responsibility of the described careers,"* and *"This unit helped me to understand what skills are needed in the described careers."* These measures of perceived learning capture students' confidence in their career knowledge, which may in turn be more predictive of career choice than actual knowledge. While the general knowledge measures are short, and therefore practical for administration, they miss the specific careers present in instruments like the CVMT. These approaches can be combined by asking students to report how much they think they know about a sampling of specific careers. Consequently, student breadth of knowledge (and confidence in that knowledge) could be assessed in a targeted and practical way.

Rationale for a New Instrument

In this manuscript, we describe a new measurement instrument, SASS, to assess high school STEM career attitudes. Our instrument builds on the SCCT framework. As described above, previous instruments focused on self-efficacy, outcome expectations, interests, and goals, and did not include items to measure career awareness or perceptions of scientists. One instrument, designed to measure the impact of career-focused high school bioinformatics curricula, includes career awareness with the SCCT constructs of self-efficacy and interests (Kovarik et al., 2013). However, this instrument did not include outcome expectations, a central factor in SCCT. We aimed to develop an instrument that integrates the measures of SCCT with career awareness and perceptions of scientists so we could measure the impact of our classroom intervention on STEM career awareness and perceptions. In this paper, we detail the development and psychometric validation of this new instrument. To our knowledge, this is the first instrument that has integrated career awareness and perceptions of scientists with SCCT.

Methods

Instrument Development

We developed the SASS instrument over the course of two years. Items for SASS were drawn from the SCCT instrument for engineering majors (kindly provided upon request by R.W. Lent; Lent et al., 2005), which included the following latent variables: self-efficacy, outcome expectations, interests, goals, social supports, and social barriers. We adapted items from these latent variables into high school appropriate language. Since we were developing a general STEM version of the SCCT

questionnaire, the items were also adjusted to reflect STEM issues more broadly (Table 1 and Table S1). Throughout this manuscript, we use the following abbreviations to refer to the latent variables in the SASS instrument: SE (self-efficacy), OE (outcome expectations), IN (interests), CA (career awareness), and SP (perceptions of scientists).

We developed item content and wording for SASS using an iterative procedure, encompassing four pilot versions with over 1,000 responses. Items were reworded, dropped, or added based on expert opinions, literature and theory, and item performance in exploratory factor analysis (EFA). Following factor analysis, items were considered for rewording or dropping based on three main considerations: 1) coefficients below 0.3 on any expected factor, 2) cross-loading of the item on multiple factors, and 3) loading of the item on an unexpected factor, in contradiction to the a priori model (Costello & Osborne, 2005). During development, items related to the SCCT constructs of social barriers and goals were dropped. Items derived from the social supports variable were strongly correlated with outcome expectation items, which we expected based upon item wording and theoretical considerations. Table S1 demonstrates the evolution of several questions from the engineering SCCT (Lent et al., 2005) to our final instrument.

The 40 items in the final version of the SASS questionnaire are listed in Table 1. All questions were grouped by predicted construct, rather than randomized, in order to maintain the engineering SCCT survey layout (Lent et al., 2005). Items in Table 1 are listed in the order they were presented to students. The original engineering SCCT was scored on a 10-point or 5-point Likert scale depending on the item. SASS is scored on a 6-point Likert scale for all items. In addition to the SCCT items, we included several demographic (e.g., race, gender, year in school, and school name) and background

Table 1. SASS instrument item wording

Item		Scale (6-point Likert)
Self-Efficacy Experience (SEE)		
	I am confident in my ability to succeed in a career that requires:	No confidence-Complete confidence
Q3_1	Understanding or interpreting graphs	
Q3_2	Knowing science and using math	
Q3_3*	Informing people about topics in science and/or medicine	
Q3_4*	Figuring out how things work	
Q3_5*	Doing lab work or experiments	
Q3_6*	Talking about science	
Self-Efficacy Academic (SEA)		
	I am confident in my ability to earn a B or better in:	No confidence-Complete confidence
Q9_1	Biology classes	
Q9_2*	Chemistry classes	
Q9_3*	Physics classes	
Q9_4*	Math classes	
	How much do you agree or disagree with these statements:	Completely disagree-Totally agree
Q10_1	I feel prepared to take college courses that use science knowledge	
Q10_2*	I feel prepared to take college courses that use math skills	
Outcome Expectations (OE)		
	Having a career that uses science, technology, engineering, and/or math (STEM) will allow me to:	Completely disagree-Totally agree
Q12_1*	Earn a good salary	
Q12_2*	Get respect from other people	
Q12_3	Do exciting work	
Q12_4	Increase my self-esteem	
Q12_5*	Have a career that my family values	
Q12_6*	Get a job that is in high demand	
	If I choose to pursue a science, technology, engineering, and/or math (STEM) related career, I will:	Not likely-Very likely
Q13_1	Get encouragement from my friends	
Q13_2	Be supported in this decision by my family members	
Q13_3	Feel that my friends and/or family are proud	
Q13_4	Feel that I “fit in” with other people in this field	
Interest (IN)		
	Rate your interest in:	Low interest-High interest
Q14_1*	Thinking about topics that relate to my health	
Q14_2*	Hearing about how researchers solve health problems	
Q14_3*	Helping people understand the importance of science in their daily lives	
Q14_4*	Watching videos or listening to podcasts that are about science, math, or health topics	
Q14_5	Thinking about “hot topics” in any of the following: science, technology, engineering, and/or math (STEM)	
Perceptions of Scientists (SP)		
	How much do you agree or disagree with this statement: A scientist is a person who...	Completely disagree-Totally agree
Q29_1*	Has an unhappy marriage	
Q29_2*	Only cares about their work	
Q29_3*	Is lonely	

(table continues)

Career Awareness (CA)		
	How much do you know about these careers?	Little or nothing-Quite a bit
Q16_1	Doctor	
Q16_2	Animal care technician	
Q16_3*	Environmental regulator	
Q16_4	Architect	
Q16_5	Computer programmer	
Q16_6*	Research scientist	
Q16_7*	Data scientist	
Q16_8*	Clinical trial recruiter	
Q16_9	Pharmacist	
Q16_10	Toxicologist	

*Item retained in reduced six factor model.

(e.g., grades in school, intention to go to college, intention to pursue a STEM career, science courses taken, and science activities outside of school) questions in the survey (Table S2). These questions were collected to help understand the environmental factors impacting students' self-efficacy, outcome expectations, and interests in future studies (as discussed in: Lent et al., 2000).

The engineering SCCT self-efficacy items refer to confidence about future academic performance (e.g., ability to obtain a B or better in engineering classes) and confidence in the ability to cope with social barriers (e.g., ability to continue in the engineering major without support from professors or not feeling well-liked). We adapted the academic performance items to be appropriate for the high school audience, but we dropped the questions about barrier coping. Instead, we included questions about confidence in students' ability to succeed in a career that requires various STEM skills (e.g., figuring out how things work or talking about science). These new items were hypothesized to be part of the SE construct.

We developed the SP construct with the goal of learning how contemporary high school students perceive scientists. As discussed above, previous research demonstrated a correlation between SCCT constructs (self-efficacy, interests, and

goals) and perceptions of scientists, supporting the inclusion of perceptions of scientists within an SCCT framework (Christidou, 2011; Wyer, 2003). The SP items in SASS were inspired by the work of Mead and Métraux (1957) and we initially included items representing both positive and negative stereotypes of scientists. In preliminary factor analyses with early versions of SASS, these items consistently factored into separate positive and negative constructs, as was observed previously (Marshall et al., 2010). To make the latent construct most coherent, we retained the items regarding the negative perceptions of scientists and dropped items related to positive stereotypes.

The STEM career awareness items ask students to self-report their knowledge of specific careers (6-point Likert scale; little or nothing - quite a bit). Careers were chosen based on relatedness to a curriculum we created in parallel to SASS (Schneider et al., 2018, 2019). For example, our curriculum introduces students to the careers of animal care technician, research scientist, data scientist, and toxicologist, all of which are included in the latent variable. During pilot testing, we evaluated different lists of STEM careers, and the EFA construct was robust to the careers included. Accordingly, the CA items can be tailored by the investigator to reflect their particular research interests.

Study Participants and Data Collection

All versions (pilot and final) of SASS were tested on high school students in the northeastern United States who were enrolled in a STEM class. Summary statistics of participants included in the reliability and validity testing of the final version are presented in Table S2. Survey instruments and procedures involving research study participants and data management were approved by the Institutional Review Board of Tufts University School of Medicine.

Surveys were administered through an online platform; the final version of SASS was hosted on Qualtrics. Responses for each item were requested from participants, but not required. Teachers were asked to administer the survey in class to increase the likelihood that students would take the survey seriously. To account for this, the survey included the following question: “Where are you taking this survey? a) At school in class, b) At school on my own time, c) At home, or d) Other: please specify.” A small subset of students (N = 104; 11.2%) reported taking the survey outside of class.

Socio-economic data for participating schools (as they describe, “percent economically disadvantaged”) was collected from New York (https://data.nysed.gov; downloaded on Oct 20, 2020) and Massachusetts (https://profiles.doe.mass.edu; downloaded on Oct 19, 2020) state databases (data was not obtained for students from other states). Downloaded files are available at: https://gitlab.com/emcdonough/sass-validation. In both state databases, students were deemed economically disadvantaged based on enrollment in one or more government assistance programs (e.g., food stamps, foster care, Earned Income Tax Credit, Safety Net Assistance Program, etc.).

Data Analysis

All data analysis was performed using R v 3.5.2 (R Core Team, 2013). The code and data used to generate results for this paper are available at: https://gitlab.com/emcdonough/sass-validation. Participants with any missing data (*n* = 44; 4.4%) were removed from the dataset. Of these, 33 (3.3%) withdrew from the study as indicated by failing to finish the survey and 10 (1.0%) participants were missing responses to up to four items. One participant took over 7 days to complete the survey, and only completed the second half. Two methods, similar to those reported previously (Adams et al., 2006; Semsar et al., 2011), were used to assess whether students

provided valid responses. First, students exhibiting string response behavior (e.g., answering 95% 5s or 1s; *n* = 3) were removed from the dataset. Second, the length of time a student took to fill in the survey was automatically recorded by Qualtrics, and this was used to remove students who likely responded to questions without reading them. Based on data from three researchers, an average of 3 seconds per item was determined as the quickest a respondent could fill out the SASS and associated demographic questions and comprehend what they were reading. Therefore, students who completed the survey more quickly than 198 seconds (*n* = 10) were removed from the dataset. The final sample included 932 high school students.

Validity Testing

Content validity of all items was assessed by six experts in education research, qualitative and quantitative data analysis, psychometric research, and laboratory science. Items were checked for clarity and relevance to the high school audience. Any items chosen for rewording were reevaluated by the expert group.

The internal structure of the survey was identified and confirmed using factor analysis to determine construct validity. Due to the addition of our two new latent variables, SP and CA, we decided to first perform an EFA. Following the EFA, the factor structure was confirmed using confirmatory factor analysis (CFA). The sample was randomly split to assign 400 participants to the EFA sample and 532 to the CFA sample. These sample sizes are considered robust for factor analysis (Guadagnoli & Velicer, 1988; Rouquette & Falissard, 2011).

Each sample was first tested for factorability using three methods as recommended by other authors (Field, 2013; Knekta et al., 2019). First, visual examination of an inter-item correlation plot (generated using the “corrplot” package for

R, version 0.84; Wei & Simko, 2017) was conducted to ensure that the majority of correlations were between 0.2 and 0.8. Second, the Kaiser-Meyer-Olkin (*KMO*) measure of sampling adequacy was calculated using the R package 'psych' (version 2.0.9; Kaiser, 1974; Revelle, 2020). *KMO* values above 0.8 were considered *meritorious* while values above 0.9 were considered *marvelous* (Field, 2013). Finally, Bartlett's test of sphericity was measured to ensure relatedness between items (Bartlett & Fowler, 1937). Bartlett's test measures whether the item correlation matrix is significantly different from an identity matrix. While a statistically significant result is generally obtained, especially with a large sample size, the lack of significance is an indication of a factorability problem (Field, 2013). Normality was examined using the 'psych' package (Revelle, 2020). Univariate normality or non-normality was determined by examining item skew and kurtosis. Item skew and kurtosis below |2.0| were considered normal (Kim, 2013; Knekta et al., 2019). Multivariate normality or non-normality was determined by visual examination of a QQ plot.

The 'psych' package was used to compute the EFA (Revelle, 2020). Considering the non-normal nature of the data (see Results) the principal axis factoring method was used for factor extraction in the EFA with an oblique rotation (oblimin; Costello & Osborne, 2005; Knekta et al., 2019; Watson, 2017). Theoretical considerations, examination of the scree plot, and parallel analysis were used to determine the proper number of factors to extract.

The R package 'lavaan' (version 0.6-7; Rosseel, 2012) was used to conduct the CFAs with the robust maximum likelihood estimator due to the continuous nature and slight non-normality of the data (Knekta et al., 2019). We used the root-mean-square error of approximation (*RMSEA*; absolute fit index; adequate fit < 0.1), the comparative

fit index (*CFI*; incremental fit index; adequate fit > 0.9), and the Tucker-Lewis index (*TLI*; incremental fit index; adequate fit > 0.9) to evaluate model fit (Browne & Cudeck, 1992; Hu & Bentler, 1999; Knekta et al., 2019; Oh et al., 2013; Tucker & Lewis, 1973). The chi-square value was also calculated. However, chi-squares are quite sensitive to sample size, and a large sample size can result in rejection of an adequate model (Knekta et al., 2019). We calculated the χ^2/df and considered values less than 2 as adequate (Oh et al., 2013). Marsh and colleagues caution that strict adherence to goodness-of-fit index cut-offs can result in the rejection of valid models (Marsh et al., 2004). Therefore, while model fit indices were assessed, they were considered in conjunction with theoretical interpretability when presenting a final model.

To assess whether the factor structure was valid for different populations, we performed a series of measurement invariance tests. For these analyses, students who responded “Prefer not to answer” or did not reply to the specific demographic items were excluded from analysis. We defined ‘minority’ as any non-white student, including those who reported belonging to multiple races. School SES was estimated using state reporting of “percent economically disadvantaged.” These values were z-standardized. Schools with z-scores greater than 0 were considered low SES, while those with z-scores less than 0 were considered high SES.

Criterion-related validity was assessed through a series of correlations of latent variables from SASS with external variables. Student latent variable scores were obtained by averaging item responses for each construct. Spearman correlations of the latent variable scores with self-reported STEM GPA were calculated. Students reported individual grades for math, biology, chemistry, and physics, which were converted to standard GPA scores (1 = D, 2 = C, 3 = B, and 4 = A) and averaged to arrive at the

STEM GPA score. Additionally, Spearman correlations between school SES and the latent variable scores were assessed. Prior to calculating correlations, distributions were examined for normality using QQ plots, and outliers surpassing 3 standard deviations were excluded. This resulted in the removal of 5 outliers for STEM GPA, and no outliers for school SES.

Reliability Testing

Internal consistency of each scale was assessed using two methods. First, Cronbach's alpha was determined using the ‘psych’ package (Revelle, 2020), as this is the statistic most widely accepted for reliability measures. However, several researchers caution against the reliance on Chronbach’s alpha for assessing instrument reliability due to several limitations of the method (Boateng et al., 2018; Knekta et al., 2019; Peters, 2014). Therefore, the omega coefficient, which has been suggested as a replacement for alpha (McNeish, 2018; Peters, 2014), was calculated using the ‘semTools’ package (version 0.5-3; Jorgensen et al., 2020).

Test-retest reliability of SASS was examined using a subset of 62 student responses collected from two schools. Teachers were asked to administer the survey twice in the classroom 7 to 10 days apart, however the time difference in administration ranged from 4 to 12 days. Bland-Altman plots were created using the ‘blandr’ package (version 0.5.1; Datta, 2017) to visually examine the agreement between the first and second measurement of the SASS (Bland & Altman, 1986). Additionally, we examined test-retest reliability by calculating type 3 intraclass correlation coefficients (*ICC(3,1)*; two-way, mixed model, single measures; Shrout & Fleiss, 1979) for each of the 40 items and the six factor scores. The *ICC(3,1)* values and confidence intervals were calculated using the ‘psych’ package. This method is recommended for determining the agreement

between two measurements taken using a self-report survey with the same subjects under the same conditions (Koo & Li, 2016; Portney & Watkins, 2015). Following the recommendations of Koo and Li (2016), we considered the following rating scale for *ICC(3,1)* values: *poor* - less than 0.5, *moderate* - 0.5 to 0.75, *good* - 0.75 to 0.90, and *excellent* - greater than 0.9.

Results

Exploratory Factor Analysis

To determine the dimensionality of the new instrument, we performed an EFA using responses from 400 students. The data exhibited good factorability (*KMO* = 0.9; Bartlett's *K-squared* = 357.55, *df* = 39, $p < 10^{-15}$) and the inter-item correlation matrix exhibited several correlations above 0.3 (Figure S1). All items had skewness and kurtosis below |2.0|, except for two items: “If I choose to pursue a science, technology, engineering, and/or math (STEM) related career, I will be supported in this decision by my family members” (kurtosis = 2.9) and “A scientist is a person who has an unhappy marriage” (kurtosis = 3.2). Examination of the QQ plot suggested possible non-normality of the data. With consideration of the items with non-normal kurtosis and the results from the QQ plot, we proceeded with the principal axis factoring extraction method, which is robust to slightly non-normal data (Knekta et al., 2019).

In addition to our two new latent variables (SP and CA), SASS included items from four constructs from the engineering SCCT (with modifications and additions relevant to high school students and STEM): self-efficacy, outcome expectations, interests, and social supports (Lent et al., 2005). As noted in the Methods, during instrument development, we observed that items regarding expectations of social

support were highly correlated with the OE items, suggesting those items could be combined with OE to form a five factor solution (corresponding to SE, OE, IN, SP, and CA). However, examination of the scree plot suggested six to seven factors based on the inflection point (Figure S2), and parallel analysis suggested an optimal solution that specified eight factors. Five factors exhibited eigenvalues above 1. Considering these results, and our theoretical expectations (we did not expect to identify more than six factors), we performed two EFAs, one with five factors and one with six factors. Theoretical considerations based on a priori knowledge, along with examination of the pattern matrices, were used to evaluate and compare model fit for the five factor and six factor models (Table 2, Table S3, and Figure S3).

The five factor solution explained 50% of the variance in the model. Each of the factors had a sum of squared loadings above 2, indicating that they were contributing positively to the model. All cross factor loadings were below 0.30, except for three items. These were equal to or below 0.40 ("I am confident in my ability to succeed in a career that requires knowing science and using math", "I feel prepared to take college courses that use science knowledge", and "Having a career that uses science, technology, engineering, and/or math (STEM) will allow me to do exciting work"). All three of these items cross-loaded on the a priori expected IN factor. Furthermore, four items which we expected to load on the SE factor instead loaded on IN above 0.50 ("I am confident in my ability to succeed in a career that requires: -informing people about topics in science and/or medicine, -figuring out how things work, -doing lab work or experiments, and -talking about science"). Pattern coefficients were above 0.50 for 30 of the 40 items. All items in the IN factor loaded poorly, with pattern coefficients below 0.45. One item ("Rate your interest in: thinking about topics that relate to my health") did

Table 2. Item characteristics for six factor EFA and CFA

		EFA Results (*n* = 400)									CFA Results (*n* = 532)					
Category	**Item**	**PA5**	**PA4**	**PA1**	**PA6**	**PA3**	**PA2**	***h2***	***u2***	***com***	**SEE**	**SEA**	**OE**	**IN**	**SP**	**CA**
SEE	Q3_1	**0.32**	**0.36**	0.05	-0.07	0.03	0.13	0.35	0.65	2.39	0.58					
SEE	Q3_2	**0.44**	**0.53**	0.08	-0.14	0.03	0.08	0.65	0.35	2.21	0.66					
SEE	Q3_3*	**0.64**	0.00	0.06	0.21	-0.09	0.02	0.63	0.37	1.27	0.73					
SEE	Q3_4*	**0.66**	0.08	0.06	-0.02	-0.05	-0.01	0.51	0.49	1.06	0.68					
SEE	Q3_5*	**0.64**	0.01	0.01	-0.05	-0.08	0.03	0.43	0.57	1.05	0.69					
SEE	Q3_6*	**0.73**	0.01	0.05	0.18	0.00	0.04	0.73	0.27	1.13	0.79					
SEA	Q9_1	0.11	**0.49**	0.08	0.29	-0.03	-0.09	0.50	0.50	1.91		0.71				
SEA	Q9_2*	0.00	**0.64**	-0.01	0.24	-0.07	-0.03	0.53	0.47	1.29		0.78				
SEA	Q9_3*	-0.05	**0.75**	-0.01	0.17	0.03	-0.03	0.61	0.39	1.12		0.76				
SEA	Q9_4*	-0.04	**0.72**	0.11	-0.16	-0.01	0.01	0.55	0.45	1.15		0.60				
SEA	Q10_1	**0.31**	**0.31**	0.09	0.22	-0.06	0.11	0.54	0.46	3.37		0.67				
SEA	Q10_2*	0.13	**0.61**	0.06	-0.08	-0.02	0.13	0.50	0.50	1.24		0.63				
OE	Q12_1*	0.09	0.10	**0.61**	-0.05	0.06	0.04	0.46	0.54	1.14			0.67			
OE	Q12_2*	0.08	-0.02	**0.67**	0.02	0.12	0.01	0.47	0.53	1.10			0.65			
OE	Q12_3	0.24	-0.06	**0.45**	0.22	-0.05	0.00	0.49	0.51	2.11			0.66			
OE	Q12_4	0.17	-0.08	**0.53**	0.13	-0.02	-0.03	0.43	0.57	1.41			0.69			
OE	Q12_5*	0.07	-0.07	**0.79**	-0.02	0.04	-0.01	0.61	0.39	1.04			0.73			
OE	Q12_6*	0.10	-0.04	**0.77**	-0.06	0.05	0.00	0.58	0.42	1.06			0.71			
OE	Q13_1	-0.09	0.08	**0.52**	0.16	-0.09	0.10	0.45	0.55	1.46			0.68			
OE	Q13_2	-0.19	0.11	**0.70**	0.10	-0.10	-0.05	0.57	0.43	1.30			0.66			
OE	Q13_3	-0.15	0.08	**0.76**	-0.01	-0.12	0.04	0.62	0.38	1.16			0.70			
OE	Q13_4	0.13	0.10	**0.51**	0.10	-0.05	0.07	0.50	0.50	1.36			0.62			
IN	Q14_1*	-0.05	0.06	0.11	**0.59**	-0.03	0.00	0.43	0.57	1.11				0.63		
IN	Q14_2*	0.06	0.01	0.06	**0.67**	-0.05	0.04	0.57	0.43	1.05				0.77		
IN	Q14_3*	0.09	-0.02	0.10	**0.65**	-0.02	0.17	0.67	0.33	1.23				0.81		
IN	Q14_4*	0.13	0.05	-0.03	**0.60**	0.06	0.10	0.48	0.52	1.20				0.75		
IN	Q14_5	0.10	0.16	0.05	**0.57**	0.00	0.12	0.58	0.42	1.33				0.79		
SP	Q29_1*	0.05	-0.06	-0.01	0.00	**0.86**	-0.02	0.75	0.25	1.02					0.87	
SP	Q29_2*	-0.07	0.04	0.04	0.04	**0.82**	0.01	0.66	0.34	1.03					0.78	
SP	Q29_3*	-0.03	0.03	-0.01	0.00	**0.90**	0.02	0.82	0.18	1.01					0.90	
CA	Q16_1	0.12	-0.08	0.18	0.11	-0.04	**0.40**	0.35	0.65	1.89						0.42
CA	Q16_2	0.05	-0.16	-0.05	0.01	-0.04	**0.54**	0.30	0.70	1.21						0.47
CA	Q16_3*	0.01	-0.11	-0.01	0.02	0.03	**0.66**	0.43	0.57	1.06						0.55
CA	Q16_4	0.03	0.05	-0.01	-0.08	0.05	**0.55**	0.30	0.70	1.09						0.54
CA	Q16_5	-0.01	0.11	0.02	-0.12	0.03	**0.66**	0.42	0.58	1.13						0.59
CA	Q16_6*	0.03	0.01	0.01	0.08	-0.05	**0.72**	0.61	0.39	1.04						0.82
CA	Q16_7*	-0.08	0.09	-0.01	0.00	-0.02	**0.83**	0.68	0.32	1.04						0.80
CA	Q16_8*	-0.07	-0.03	-0.02	0.08	-0.02	**0.66**	0.44	0.56	1.06						0.68
CA	Q16_9	0.06	-0.05	0.11	0.11	0.04	**0.58**	0.46	0.54	1.19						0.62
CA	Q16_10	0.05	-0.09	-0.04	0.04	0.03	**0.66**	0.45	0.55	1.07						0.63

See Table 1 for item wording. Asterisks (*) indicate items retained in the reduced six factor model. EFA pattern coefficients above 0.3 are presented in bold. For the CFA results, all factor loadings were statistically significant at $p < 10^{-3}$. Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, CA = career awareness, PA[1-6] = principal axis factor, *h2* = communality, *u2* = unique variance, and *com* = Hoffman's index of complexity.

not load above 0.27 on any factor. In sum, these results suggested that the addition of a sixth factor to account for the four SE items loading with IN may improve the model.

The six factor model explained a greater percentage of the variance (53%) than the five factor model and all factors had a sum of squared loadings above 2. The four SE items that unexpectedly loaded with the IN factor in the five factor model comprised a new factor when a six factor model was examined. Based on examination of the item wording, we chose to call this new factor SE-experience (SEE) and categorized the remaining SE items as SE-academic (SEA). Three items were cross-loaded in this six factor model ("I am confident in my ability to succeed in a career that requires: -understanding or interpreting graphs, -knowing science and using math," and "I feel prepared to take college courses that use science knowledge") and they had roughly equivalent pattern coefficients for the a priori SEA factor and SEE. We chose to classify the first two items as SEE and the third as SEA based on our evaluation of the item wording. All items had pattern coefficients above 0.50, except for five, which were above 0.30. Of the 40 items in the instrument, 25 exhibited pattern coefficients above 0.60.

Confirmatory Factor Analysis

Following EFA, CFA was performed using an independent sample of 532 student responses. The data exhibited good factorability (*KMO* = 0.9; Bartlett's *K-squared* = 602.23, *df* = 39, $p < 10^{-15}$) and the inter-item correlation matrix exhibited several correlations above 0.30 (Figure S4). All items had skewness and kurtosis below |2.0|, except for three items: "If I choose to pursue a science, technology, engineering, and/or math (STEM) related career, I will be supported in this decision by my family members" (kurtosis = 2.2), "A scientist is a person who has an unhappy marriage" (kurtosis = 2.5),

and “A scientist is a person who is lonely” (kurtosis = 2.0). As seen with the EFA sample, examination of the QQ plot suggested slight non-normality of the data. With consideration of the items with non-normal kurtosis and the results from the QQ plot, we proceeded with the robust maximum likelihood estimation method due to the continuous nature and the slight non-normality of the data (Knekta et al., 2019). For all models tested, all item estimates were significant ($p < 10^{-3}$).

We began by constructing five factor and six factor models of the data (Figure 1, Table 2, and Table S4). A single factor model was also calculated and is provided for reference (Table S6). Fit indices for the five factor and six factor model were similar, with *CFI* and *TLI* values between 0.7 and 0.8, and *RMSEA* of 0.08 (Table 3). Neither of the models exhibited a χ^2/df less than 2 (the suggested cut-off), but this may be because the measure is inaccurate for larger sample sizes (Knekta et al., 2019). While the *CFI* and *TLI* values were below the cut-off for an adequate fit (0.9), the *RMSEA* was below 0.1, indicating an adequate fit by that measure (Browne & Cudeck, 1992; Hu & Bentler, 1999; Knekta et al., 2019; Oh et al., 2013; Tucker & Lewis, 1973). For both models, all item loadings were between 0.50 and 0.90, with two exceptions: CA items regarding doctor (five factor: 0.41; six factor: 0.42) and animal care technician (five factor: 0.47; six factor: 0.47). Factor loadings were minimally changed for most items when comparing the five and six factor models, with the exception of all items in the SE, SEE, and SEA factors, which generally exhibited higher loading scores in the six factor model (average difference in item loading = 0.06). In sum, the CFA and EFA results both suggest that the six factor model is a better fit for the data.

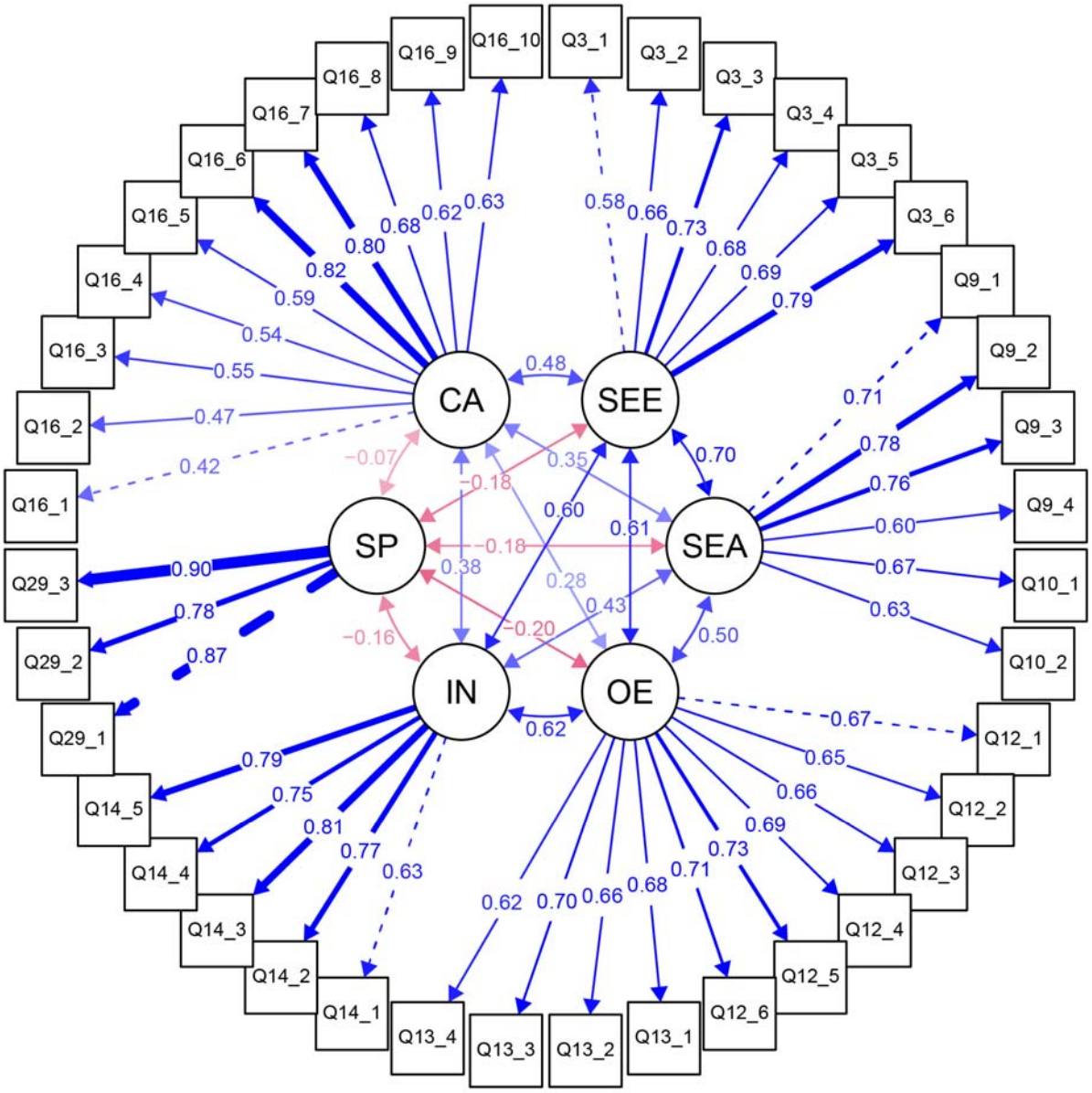


Figure 1. Confirmatory factor analysis item loadings and factor correlations

Results from the six factor CFA model are shown. Factors are represented by circles, while survey items are represented by squares (see Table 1 for item descriptions). The one-directional arrows from the factors to the items represent the standardized factor loadings; the arrows are weighted by the strength of the factor loading and the dashed arrows indicate which item estimates were fixed to 1 in the model. The bi-directional arrows between factors represent the correlations between the factors. Blue arrows indicate a positive loading or correlation. Red arrows indicate a negative loading or correlation. All item estimates were significant at $p < 10^{-3}$. Latent variable correlations were significant at $p < 10^{-2}$, except for SP with CA, which was $p = 0.20$.

Table 3. CFA fit indices

Model	*n*	χ^2	*df*	χ^2/df	*CFI*	*TLI*	*RMSEA*	Lower	Upper
One factor	532	6999	740	9.5	0.45	0.42	0.13	0.12	0.13
Five factor	532	3506	730	4.8	0.76	0.75	0.08	0.08	0.09
Six factor	532	3231	725	4.5	0.79	0.77	0.08	0.08	0.08
Men	248	1994	725	2.8	0.77	0.75	0.08	0.08	0.09
Women	269	1990	725	2.7	0.79	0.77	0.08	0.08	0.08
White	257	1956	725	2.7	0.79	0.78	0.08	0.08	0.09
Minority	191	1967	725	2.7	0.71	0.69	0.09	0.09	0.10
High SES	184	1979	725	2.7	0.72	0.70	0.10	0.09	0.10
Low SES	265	2000	725	2.8	0.79	0.77	0.08	0.08	0.09
Reduced six factor	532	678	215	3.2	0.92	0.91	0.06	0.06	0.07

For *CFI*, *TLI*, *RMSEA*, lower, and upper, robust measures were used. Lower and upper refer to the bounds of the 90 percent confidence interval estimate for RMSEA. Abbreviations: *n* = number of students, χ^2 = chi-square, *df* = degrees of freedom, *CFI* = comparative fit index, *TLI* = Tucker-Lewis index, and *RMSEA* = root-mean-square error of approximation.

Correlations between latent factors for the six factor model are presented in Table 4. The latent factor correlations range from -0.2 (SP with SEA, SEE, OE, and IN) to 0.7 (SEE with SEA). In general, the SCCT latent variables (SEE, SEA, OE, and IN) were highly correlated with each other. The CA factor exhibited moderate correlation with the SCCT factors, and SP was weakly correlated with these factors.

Table 4. CFA six factor model latent variable correlations

	SEE	SEA	OE	IN	SP
SEA	0.70*	-			
OE	0.61*	0.50*	-		
IN	0.60*	0.43*	0.62*	-	
SP	-0.18*	-0.18*	-0.20*	-0.16*	-
CA	0.48*	0.35*	0.28*	0.38*	-0.07

* $p < 10^{-2}$. Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

With the goal of constructing an optimally performing CFA model, we selected the best performing items from the six factor EFA. This was achieved by first sorting items within a latent variable by their Hoffman complexity scores, which measures the average number of factors necessary to explain each item and is included in the output from the EFA (Table 2). An ideal Hoffman's complexity score is 1, indicating that one factor is required to explain the item (Knekta et al., 2019). For each latent variable, we retained the four items with the lowest complexity score, except for SP, which only contained three items in SASS. This resulted in the retention of 22 items (Table 1 and Table 2). This 'reduced six factor model' produced *CFI* and *TLI* values greater than 0.9, and an *RMSEA* of 0.06. These results suggest that the reduced six factor model is a better fit for the data than the five or six factor models.

While developing the reduced six factor model, nearly half of the instrument's original items were discarded. Examination of the remaining 22 items revealed a more narrow instrument than intended. For example, the optimal SEA construct retained items about math, physics, and chemistry (all math heavy courses), and dropped items pertaining to science and biology. Items retained in the OE construct related to how one is perceived by others (e.g., "Get respect from other people"), and dropped items related to self-judgement (e.g., "Increase my self-esteem). Overall, the limited scope of the items included in the reduced six factor model poorly reflected the breadth of the latent variables we aimed to measure. All remaining analyses presented in this manuscript were conducted using the six factor model, with retention of all 40 items.

Confirmation of an instrument's structural validity in subgroups is often used to support construct validity (e.g., Unfried et al., 2015). Therefore, we performed CFAs with subsamples of the data, split by gender, race, and SES. Models for all subsamples

performed similarly (Table 3; *CFI* and *TLI* between 0.7 and 0.8; *RMSEA* between 0.08 and 0.10). Consistent with their smaller sample sizes, these subsample models exhibited χ^2/df closer to 2 than seen with the full sample set. These results indicate that the six factor model structure is acceptable and stable across race, gender, and SES in our sample.

Criterion Validity

One method of determining instrument validity is to test the relationship of each latent variable within the instrument to an independent criterion. We compared self-reported STEM GPA to SEE scores and SEA scores, which measure students' confidence in using STEM-related skills or their confidence in their ability to obtain a B or better in STEM classes, respectively (Figure 2A-B). Considering the shared relationship of STEM GPA, SEE, and SEA to academic performance, we hypothesized that both SEE and SEA scores would increase with students' self-reported STEM GPAs. As expected, STEM GPA was positively correlated with SEE and SEA scores in our sample (SEE: $rho = 0.33$, $p < 10^{-14}$; SEA: $rho = 0.62$, $p < 10^{-15}$). A meta-analysis of 121 independent samples found a correlation between academic achievement and academic interest (Schiefele et al., 1992). Therefore, we examined the relationship between student STEM GPA and IN scores (Figure 2C). As expected, a statistically significant correlation was found ($rho = 0.21$, $p < 10^{-5}$). Previous research suggested a relationship between SES and self-efficacy (Navarro et al., 2007), as well as SES and outcome expectations (Turner et al., 2019). Therefore, we examined the correlation between SES and SEE, SEA, and OE scores (Figure 2D-F). Our results indicated that a lower percent economically disadvantaged (higher SES) was associated with higher SEE ($rho = -0.12$, $p = 10^{-2}$), SEA ($rho = -0.18$, $p < 10^{-3}$), and OE ($rho = -0.14$, $p < 10^{-2}$)

scores. For completeness, we examined the correlation of STEM GPA and SES with the remaining latent variables (Figure S5). No correlations were significant using a $p < 10^{-2}$ cutoff, except STEM GPA with OE score (*rho* = 0.24, $p < 10^{-7}$). Taken together, these results suggest that our new instrument has good criterion validity.

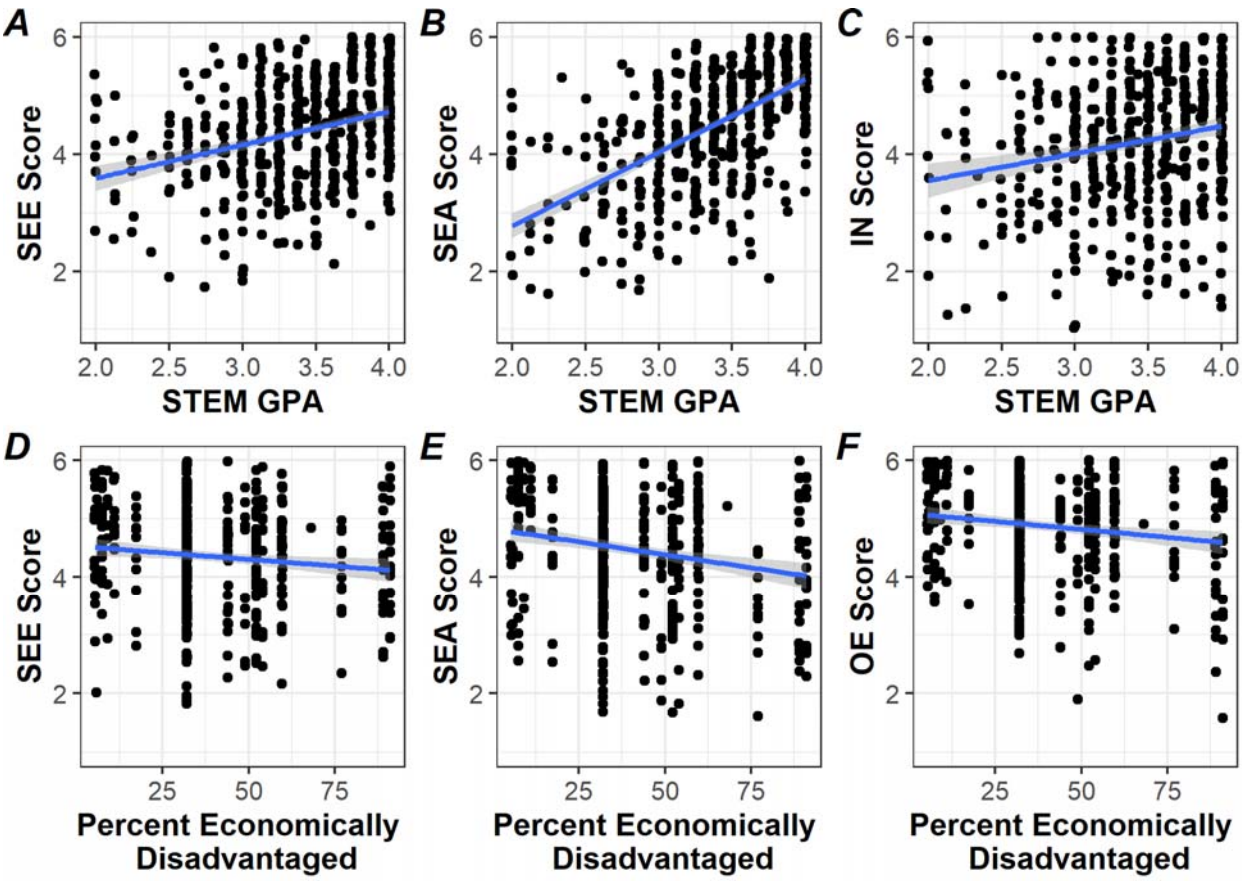


Figure 2. Criterion-related validity testing

Scatter plots illustrating the relationship of factor scores and external variables are shown with linear regression lines. Spearman's correlations and *p* values were also calculated for SASS factors with STEM GPA: A) SEE score (*rho* = 0.33, $p < 10^{-14}$); B) SEA score (*rho* = 0.62, $p < 10^{-15}$); C) IN score (*rho* = 0.21, $p < 10^{-5}$); and for SASS factors with school percent economically disadvantaged: D) SEE score (*rho* = -0.12, $p = 10^{-2}$); E) SEA score (*rho* = -0.18, $p < 10^{-3}$); and F) OE score (*rho* = -0.14, $p < 10^{-2}$). A high percentage of economically disadvantaged students corresponds with low SES for the school. Because the SES data is collected at the school level, the values are clustered at discrete values on the horizontal axis. Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

Reliability

Internal consistency of an instrument is a measure of how well each of the items in a latent variable relate to each other. We evaluated the internal consistency of the instrument by measuring Chronbach's alpha, the "gold standard," and omega for each of the six constructs. Ideally the alpha and omega values should be greater than 0.70. However, values that are greater than 0.95 would be indicative of high redundancy in the instrument items (Tavakol & Dennick, 2011). For SASS, alpha values were 0.8 for SEE and SEA, and 0.9 for OE, IN, SP, and CA. Omega values were identical to alpha except for SEA which was 0.9. Therefore, the instrument demonstrated *good* to *excellent* internal consistency.

While many SCCT instruments are used to take a snapshot of participant views, we planned to use SASS for evaluating the impact of a five day classroom intervention on student attitudes related to STEM careers. Therefore, it was important to assess the stability of the instrument over time. With a subset of student responses, we used *ICC(3, 1)* to assess the test-retest reliability of SASS (Table S7). Students in the sample subset took the survey an average of nine days apart. *ICC(3,1)* values were *good* for SEA (0.88), *moderate* for SEE, OE, IN, and SP (0.69, 0.53, 0.57, and 0.72, respectively), and *poor* for CA (0.47). Bland-Altman plots are presented in Figure S6.

Discussion

Validation and Reliability of SASS

With the goal of understanding what factors influence career choice in high school students, we developed the SASS questionnaire, which measures attitudes surrounding STEM careers. When used with a high school cohort, the instrument is

reliable, and has good structural and criterion validity. In developing SASS, hard cut-offs for the goodness-of-fit indices were carefully considered in conjunction with theory and content validity when determining model fit. Although the goodness-of-fit measures obtained for the CFA models (χ^2/df, *CFI*, *TLI*, and *RMSEA*) did not universally exceed standard cut-offs, we note that these cut-offs are meant to be taken as a 'rule of thumb,' and not applied without theoretical consideration (Marsh et al., 2004). Our sample included students from a diverse array of schools, and it is possible this high heterogeneity did not allow for superlative goodness-of-fit measures. Previous SCCT survey instruments, which exhibited goodness-of-fit metrics that exceeded the recommended cut-offs, have been validated using a more homogenous sample (e.g., Mau et al., 2019; Rasheed Ali & Saunders, 2009; Turner et al., 2019), or using data from students of a single demographic background (e.g., Garriott et al., 2013). Regardless, the *RMSEA* values for all models examined (except the single factor model) were below 0.1, indicating an acceptable fit according to this metric (Browne & Cudeck, 1992). After examination of results from EFA and CFA, as well as consideration of item wording and our study goals, we recommend the use of a six factor model for SASS.

In an effort to produce a model with exceptional fit to the data, we reduced the items to only those that performed best in the EFA. The CFA results from this reduced six factor model, which removed all but the most highly correlated and specific items, resulted in a model that exceeded standard cut-offs for *CFI*, *TLI*, and *RMSEA*. While the 40 items included in the full questionnaire represent broad factors with strong theoretical consideration, the items retained in the reduced six factor model narrowed the scope of the instrument. In general, lower correlations between individual items within a given

factor indicate a broader conceptual factor. For example, two items regarding math would be expected to correlate more closely than would one item about math and one item regarding biology. Therefore, reducing the instrument to the items that are most highly correlated would be expected to produce excellent goodness-of-fit metrics. Researchers aiming to use an instrument with exceptional structural validity could opt to use the reduced 22-item questionnaire, but in doing so, they would lose some of the meaning in the constructs. Of particular note, the SEA construct was related more to math domains than general STEM, substantially changing the interpretation of the SEA variable.

Criterion validity estimates how well a psychometric construct agrees with an external measure or a theoretical relationship. Using the six factor model, we assessed criterion validity for the SASS instrument by examining relationships of the latent variables to self-reported STEM GPA and school SES. Consistent with previous research, we found significant positive relationships between STEM GPA and SEE, SEA, and IN scores (Lopez & Lent, 1992; Navarro et al., 2007; Schiefele et al., 1992). In examining school SES as a predictor of SEE, SEA, and OE scores, we identified significant relationships between SES and the three SCCT variables, consistent with previous research (Navarro et al., 2007; Turner et al., 2019). While the relationships between SES and these SCCT constructs were significant, they were small. This is consistent with research conducted in Australia that reports relatively small differences in career aspirations related to socio-economic status (Gore et al., 2015). Additionally, we obtained school-level SES data, which obscures differences between students within a school; student-level SES data might provide more accurate estimates of the

relationships. Taken together, the correlations of SASS variables with STEM GPA and SES demonstrate criterion validity of the instrument.

We found *good* to *excellent* internal consistency of the six latent variables using alpha and omega analyses. This was expected, especially for the SEA, OE, and IN factors that were based on reliable measures from another instrument (Lent et al., 2005). Test-retest reliability of SASS factors were *good* or *moderate* for five of the six latent variables. The factor with the best test-retest reliability was SEA, suggesting that students' belief in their ability to perform well academically in the STEM fields was consistent between the two timepoints. The CA factor displayed the least consistency between the two measurements, exhibiting *poor* test-retest reliability. One explanation for low reliability of this factor may be the assumed dependence on student confidence in how much they know, which could be impacted by their mood or emotional state on any particular day. Additionally, the test-retest reliability was assessed in a small subset of students from two college-preparatory schools and may not reflect the performance of the survey in a broader sample of schools. The *ICC* values are impacted by the variability of a sample; the more homogenous a sample, the lower the *ICC* values (Aldridge et al., 2017). Therefore, it isn’t unreasonable to expect higher *ICC* values in a larger, more heterogeneous sample, and further test-retest reliability testing could be helpful. Finally, the CA factor may be sensitive to intervening events that improve knowledge about a career or reduce confidence in that knowledge. The CA factor was designed to measure the effectiveness of interventions, so sensitivity to changing career awareness is crucial.

Identification of SCCT Variables

The 40-item six factor model robustly identifies three constructs adapted from a similar SCCT instrument for engineering majors (Lent et al., 2005): SEA, OE, and IN. With the goal of broadening the SASS SE construct, we developed items to measure self-efficacy as it relates to STEM skills. A five factor model, which combined the traditional academic and newly developed skill items into a single factor, was not as coherent in the EFA results. This suggests that perceived STEM skills and perceived academic skills reflect different underlying latent constructs. For example, a student might not feel confident in trigonometry, but might feel capable in figuring out how their bicycle works. We named the new construct *self-efficacy experience* because the skills items reflect STEM experience outside the classroom. As expected, correlations among these four latent variables (SEE, SEA, OE, and IN) were high (0.43 to 0.70), confirming the well-established relationship of the SCCT constructs (e.g., Cameron et al., 2015; Lent et al., 2005, 2018; Turner et al., 2019).

A notable SCCT latent variable missing from SASS is a measure of student goals. During instrument development, we tested versions of SASS that included items which comprised a goal setting factor. These items exhibited consistently low coefficients (<0.4), cross-loaded into multiple factors, and did not have high correlations (>0.5) with each other in EFA. After four rounds of rewording questions and piloting new versions with over 1,000 students, we decided to remove the goals measure. Unlike college students, who are thinking about a major and a potential career path, high school students may not have a clear grasp on their goals. Young adolescents may have too many interests to pin down a single goal, they may conflate interests and goals, or they may not be considering their future at all. In a meta-analysis of 41 SCCT

studies, 16 out of 16 high school or middle school oriented studies did not include a goal setting measure (Sheu & Bordon, 2017). Therefore, exclusion of a goal setting latent variable is well supported for an instrument aimed at adolescents.

Inclusion of New Variables

During educational interventions, students may change their perceptions of STEM or of the people working in STEM. This is especially important for high school students who are exploring potential career aspirations, particularly if they are envisioning themselves as scientists. With this in mind, we developed the SP latent variable, to measure student perceptions of scientists. While student held perceptions of scientists were previously linked with self-efficacy, interests, and goals (Christidou, 2011; Wyer, 2003), we found small correlations (below |0.2|) between SP and the SCCT constructs. Further work, such as path analysis, could elucidate the relationship between student SP and other factors in the SCCT framework.

The three items included in SP constituted a robust and independent factor, and they conveyed the negative perceptions that scientists are isolated, work-obsessed, and have poor marriages. These stereotypes were based on work from the 1950's (Mead & Métraux, 1957). Although the items are similar to those found in two independent and valid instruments that measure stereotypes of scientists (Krajkovich & Smith, 1982; Wyer et al., 2010), it is likely that perceptions of scientists have changed in the nearly 70 years that have passed since the initial observations of Mead and Métraux. For example, the stereotype that science is for men has already started to shift, as shown by Scholes and Stahl (2020), who reported that only two out of 45 students recorded scientists as a man, while the majority of students used ungendered language. The addition of items to reflect this change in gender stereotypes could prove valuable.

Preliminary examination of the SP item responses indicated that students were extremely consistent in their scoring of the three items. Furthermore, most students scored a 1 or 2 for all three items, indicative of a floor effect. Although these items constitute a mathematically reliable and valid measure, they may not represent modern views of scientists. In other words, the participants may rate the three items similarly because they are ridiculous statements to contemporary U.S. high school students. Rewording of the current items or addition of new items may improve the factor. Additional items could include examination of scientist stereotypes related to race, class, or gender. Similarly, while the item pertaining to an “unhappy marriage” correlated highly with the other two SP items in our sample, the language is not inclusive of all forms of romantic relationships and could be updated to “romantic partnership” or “family life.” In sum, an expanded SP measure might shed more light on student perceptions and the impact of classroom interventions aimed at increasing STEM career choice.

The decision to pursue a career in STEM requires more than just traditional SCCT factors, like social support, confidence, and interests. A student must know the career exists and have some idea of what it means to perform the job. We built the CA items to reflect student knowledge of selected STEM careers. While this factor allows us to capture self-perceived knowledge of careers, it is not an objective measure of student career knowledge, and therefore it might be influenced by their confidence (i.e., self-efficacy). For example, a very insecure student might provide a low rating in knowledge of animal care technician knowledge, even if they had a family member with that job. However, the correlations among CA and the SCCT variables were low (0.28 to 0.48), suggesting that the influence of confidence on this factor may be small. Nevertheless,

self-efficacy is a complex and powerful psychological characteristic (Bandura, 2012), and further research into how SCCT constructs relate to CA could be informative.

The careers we selected for the CA factor may not be representative of all STEM careers, but these items could potentially be tailored to an intervention by a given researcher. In support of this view, early versions of SASS included other combinations of careers and we did not observe cross-loading with other factors in the preliminary EFA results. Indeed, the CFA results indicate low correlations with other factors, suggesting that career customization may be possible. Inclusion of a career awareness variable in the SCCT framework will help researchers learn how self-reported career knowledge is impacted when performing educational interventions.

Limitations

Development and validation of psychometric instruments is difficult and the results are rarely presented without limitations. An obvious limitation of any instrument validation is the sample it was tested on; factor analysis results may be different in an independent sample. While the data for this study was obtained exclusively from schools in the northeastern United States, a wide representation of school characteristics (public, private, high SES, and low SES) was included. Additionally, initial tests in subgroups of our sample population support validity of the instrument across various demographics. Regardless, the metrics reported in this study may not generalize to dissimilar populations, such as those from other countries or other regions of the United States.

The items included in the instrument were adapted from a validated instrument and were assessed for content validity by experts in the field. However, due to limited resources and time, we did not assess how adolescents interpreted the instrument

items. Detailed interviews with the target population might provide additional qualitative insight into the consistency of student interpretation. This will likely prove especially useful in further development of the SP construct.

Finally, several limitations surrounding the data collection should be noted. First, the instrument items were presented sequentially (not randomized or shuffled). This was done in order to align with the methods of Lent and colleagues (2005), as SASS was adapted from their instrument. However, order or priming effects could arise; that is, an answer could be influenced by the question that preceded it. Additionally, the validity of data collected from self-reported surveys is hard to assess because people may be dishonest. In an effort to collect valid data, we asked teachers to administer surveys during class time and requested that the surveys were not tied to grades or attendance requirements. Future versions of the instrument could include embedded questions to act as quality control measures of the data similar to Semsar and colleagues (2011).

Conclusions

In summary, we have created a useful tool for assessing educational interventions intended to impact high school student's feelings, views, and knowledge about STEM careers. Accurate assessment of these interventions is essential in preventing attrition through the "leaky pipeline." The SASS instrument extends and modifies SCCT for a high school population with the addition of several factors. We expanded the SE construct to include STEM skills outside of the classroom (SEE), and added two new factors, CA and SP. Career knowledge and negative stereotypes are crucial characteristics in determining STEM career trajectories. To our knowledge, this is the first instrument that integrates both CA and SP into the SCCT framework.

Author Contributions

E.K.M., J.H., B.J., and K.S.S. were responsible for the study concept, design, and survey item revision. E.K.M. and J.H. collected the data. E.K.M. and K.S.S. analyzed the data. The paper was written by E.K.M. and K.S.S., with additional input from the other authors. All authors critically reviewed and approved the manuscript for publication.

Acknowledgments

The authors thank Leslie Schneider, Russel Faux, Revati Masilamani, and David Reider for consultation during the collection of pilot data and the initial instrument revision process. We thank Robert W. Lent for kindly providing a copy of the SCCT survey for engineering majors. The authors thank Nicky Case for providing creative input when naming the survey instrument. Finally, we would like to acknowledge the roles of the teachers and students for making this study possible.

Research reported in this publication was supported by a National Science Foundation Innovative Technology Experiences for Students and Teachers Grant (#1614167) and a National Institute of Health Science Education Partnership Award (#R25GM137369). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the National Science Foundation, the U.S. Department of Veterans Affairs, or the United States Government.

Competing interests

The authors declare no financial or non-financial competing interests.

Supplemental Tables

Table S1. Exemplar evolution of item wording

SCCT Construct	Engineering SCCT[a]	Pilot SASS	Final SASS
Self-efficacy	How confident are you that you could: Continue on in engineering even if you felt that, socially, the environment in these classes was not very welcoming to you.	I am confident in my ability to take science classes that interest me even if I don't "fit in" with other students.	Question dropped due to poor clustering in early exploratory factor analysis.
Outcome Expectations	Graduating with a BS degree in engineering will likely allow me to go into a field with high employment demand.	Having a STEM career will allow me to get a job that is in high demand.	Having a career that uses science, technology, engineering, and/or math (STEM) will allow me to get a job that is in high demand.
Interests	How much interest do you have in reading articles or books about engineering issues.	Rate your interest in reading about STEM issues.	Rate your interest in thinking about "hot topics" in any of the following: science, technology, engineering, and/or math (STEM).

[a](Lent et al., 2005)

Table S2. Demographics

		EFA (*n* = 400)		CFA (*n* = 532)	
Variable	**Level**	***n***	**Percent**	***n***	**Percent**
Grade	9th	78	19.5	85	16.0
	10th	89	22.3	159	29.9
	11th	95	23.8	129	24.2
	12th	136	34.0	159	29.9
	No response	2	0.5		
Gender	Female	234	58.5	269	50.6
	Male	160	40.0	248	46.6
	Other / non-binary	2	0.5	3	0.6
	Prefer not to answer	1	0.3	7	1.3
	No response	3	0.8	5	0.9
Race	American Indian or Alaskan Native	4	1.0	3	0.6
	Asian	82	20.5	74	13.9
	Black or African American	55	13.8	79	14.8
	Native Hawaiian or Other Pacific Islander	2	0.5	4	0.8
	White	181	45.3	257	48.3
	Mixed race	16	4.0	31	5.8
	Prefer not to answer	50	12.5	71	13.3
	No response	10	2.5	13	2.4
Ethnicity	Not Hispanic / Latino	291	72.8	398	74.8
	Hispanic / Latino	88	22.0	111	20.9
	Prefer not to answer	18	4.5	18	3.4
	No response	3	0.8	5	0.9

Table S3. Item loadings for 5 factor EFA

Category	Item	PA4	PA1	PA5	PA3	PA2	*h2*	*u2*	*com*
SE	Q3_1	**0.42**	0.00	0.25	0.05	0.09	0.33	0.67	1.77
SE	Q3_2	**0.61**	0.00	**0.32**	0.07	0.02	0.60	0.40	1.56
SE	Q3_3	0.05	0.04	**0.72**	-0.08	0.02	0.61	0.39	1.04
SE	Q3_4	0.17	-0.01	**0.59**	-0.01	-0.04	0.42	0.58	1.18
SE	Q3_5	0.10	-0.06	**0.56**	-0.04	-0.01	0.33	0.67	1.10
SE	Q3_6	0.08	0.01	**0.79**	0.02	0.03	0.69	0.31	1.02
SE	Q9_1	**0.45**	0.15	0.25	-0.07	-0.04	0.46	0.54	1.92
SE	Q9_2	**0.58**	0.06	0.11	-0.11	0.02	0.47	0.53	1.17
SE	Q9_3	**0.70**	0.05	0.03	-0.01	0.00	0.53	0.47	1.02
SE	Q9_4	**0.76**	0.07	-0.16	0.00	-0.02	0.57	0.43	1.11
SE	Q10_1	**0.32**	0.11	**0.41**	-0.07	0.14	0.54	0.46	2.40
SE	Q10_2	**0.65**	0.02	0.05	-0.01	0.10	0.50	0.50	1.06
OE	Q12_1	0.14	**0.58**	0.05	0.09	0.01	0.44	0.56	1.18
OE	Q12_2	0.00	**0.66**	0.09	0.14	-0.01	0.46	0.54	1.13
OE	Q12_3	-0.05	**0.48**	**0.35**	-0.05	0.01	0.49	0.51	1.89
OE	Q12_4	-0.07	**0.55**	0.24	-0.01	-0.03	0.43	0.57	1.41
OE	Q12_5	-0.03	**0.77**	0.06	0.07	-0.03	0.58	0.42	1.03
OE	Q12_6	0.01	**0.73**	0.05	0.08	-0.04	0.54	0.46	1.04
OE	Q13_1	0.05	**0.58**	-0.02	-0.11	0.13	0.45	0.55	1.19
OE	Q13_2	0.08	**0.75**	-0.14	-0.10	-0.03	0.58	0.42	1.14
OE	Q13_3	0.08	**0.77**	-0.17	-0.11	0.03	0.62	0.38	1.16
OE	Q13_4	0.11	**0.52**	0.17	-0.04	0.07	0.50	0.50	1.38
IN	Q14_1	-0.04	0.27	0.27	-0.11	0.12	0.29	0.71	2.78
IN	Q14_2	-0.08	0.23	**0.40**	-0.13	0.18	0.42	0.58	2.42
IN	Q14_3	-0.10	0.26	**0.42**	-0.10	0.29	0.55	0.45	2.79
IN	Q14_4	-0.02	0.11	**0.43**	-0.02	0.21	0.37	0.63	1.61
IN	Q14_5	0.08	0.19	**0.40**	-0.08	0.23	0.48	0.52	2.27
SP	Q29_1	-0.06	-0.01	0.05	**0.86**	-0.02	0.75	0.25	1.02
SP	Q29_2	0.03	0.06	-0.05	**0.80**	0.03	0.63	0.37	1.02
SP	Q29_3	0.03	0.00	-0.03	**0.90**	0.03	0.80	0.20	1.01
CA	Q16_1	-0.07	0.18	0.17	-0.03	**0.41**	0.35	0.65	1.84
CA	Q16_2	-0.14	-0.06	0.05	-0.03	**0.55**	0.30	0.70	1.18
CA	Q16_3	-0.10	-0.02	0.01	0.04	**0.67**	0.43	0.57	1.05
CA	Q16_4	0.08	-0.04	-0.03	0.07	**0.54**	0.29	0.71	1.11
CA	Q16_5	0.14	-0.03	-0.09	0.06	**0.64**	0.40	0.60	1.17
CA	Q16_6	0.02	0.01	0.06	-0.04	**0.74**	0.61	0.39	1.02
CA	Q16_7	0.10	-0.02	-0.09	-0.01	**0.84**	0.68	0.32	1.06
CA	Q16_8	-0.03	-0.01	-0.03	-0.03	**0.68**	0.44	0.56	1.01
CA	Q16_9	-0.04	0.11	0.11	0.04	**0.60**	0.46	0.54	1.16
CA	Q16_10	-0.07	-0.05	0.06	0.04	**0.66**	0.44	0.56	1.06

See Table 1 for item wording. EFA pattern coefficients above 0.3 are presented in bold. Abbreviations: SE = self-efficacy, OE = outcome expectations, IN = interests, SP = perceptions of scientists, CA = career awareness, PA[1-5] = principal axis factor, *h2* = communality, *u2* = unique variance, and *com* = Hoffman's index of complexity.

Table S4. Item loadings for 5 factor CFA

Item	SE	OE	IN	SP	CA
Q3_1	0.60				
Q3_2	0.73				
Q3_3	0.64				
Q3_4	0.62				
Q3_5	0.61				
Q3_6	0.68				
Q9_1	0.62				
Q9_2	0.67				
Q9_3	0.64				
Q9_4	0.52				
Q10_1	0.71				
Q10_2	0.62				
Q12_1		0.67			
Q12_2		0.65			
Q12_3		0.65			
Q12_4		0.69			
Q12_5		0.73			
Q12_6		0.71			
Q13_1		0.68			
Q13_2		0.66			
Q13_3		0.70			
Q13_4		0.61			
Q14_1			0.62		
Q14_2			0.77		
Q14_3			0.81		
Q14_4			0.75		
Q14_5			0.79		
Q29_1				0.87	
Q29_2				0.78	
Q29_3				0.90	
Q16_1					0.41
Q16_2					0.47
Q16_3					0.55
Q16_4					0.54
Q16_5					0.59
Q16_6					0.82
Q16_7					0.80
Q16_8					0.68
Q16_9					0.62
Q16_10					0.63

See Table 1 for item wording. All factor loadings were statistically significant at $p < 10^{-3}$. Abbreviations: SE = self-efficacy, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

Table S5. Item loadings for reduced six factor CFA

Item	SEE	SEA	OE	IN	SP	CA
Q3_3	0.78					
Q3_4	0.65					
Q3_5	0.68					
Q3_6	0.85					
Q9_2		0.75				
Q9_3		0.79				
Q9_4		0.68				
Q10_2		0.65				
Q12_1			0.76			
Q12_2			0.69			
Q12_5			0.74			
Q12_6			0.72			
Q14_1				0.74		
Q14_2				0.88		
Q14_3				0.76		
Q14_4				0.65		
Q29_1					0.87	
Q29_2					0.78	
Q29_3					0.90	
Q16_3						0.49
Q16_6						0.85
Q16_7						0.85
Q16_8						0.65

See Table 1 for item wording. All factor loadings were statistically significant at $p < 10^{-3}$. Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

Table S6. Item loadings for 1 factor CFA

Item	SCCT
Q3_1	0.51
Q3_2	0.62
Q3_3	0.64
Q3_4	0.56
Q3_5	0.60
Q3_6	0.68
Q9_1	0.52
Q9_2	0.54
Q9_3	0.51
Q9_4	0.37
Q10_1	0.66
Q10_2	0.51
Q12_1	0.55
Q12_2	0.50
Q12_3	0.62
Q12_4	0.60
Q12_5	0.58
Q12_6	0.61
Q13_1	0.60
Q13_2	0.53
Q13_3	0.55
Q13_4	0.62
Q14_1	0.45
Q14_2	0.55
Q14_3	0.60
Q14_4	0.59
Q14_5	0.70
Q29_1	-0.20
Q29_2	-0.20
Q29_3	-0.21
Q16_1	0.38
Q16_2	0.21
Q16_3	0.24
Q16_4	0.24
Q16_5	0.36
Q16_6	0.51
Q16_7	0.43
Q16_8	0.32
Q16_9	0.36
Q16_10	0.30

See Table 1 for item wording. All factor loadings were statistically significant at $p < 10^{-3}$.

Table S7. Intraclass correlation coefficients

Variable	*ICC(3,1)*	Lower Bound	Upper Bound
Q3_1	0.75	0.64	0.83
Q3_2	0.75	0.64	0.83
Q3_3	0.60	0.44	0.72
Q3_4	0.59	0.43	0.71
Q3_5	0.37	0.18	0.54
Q3_6	0.48	0.30	0.63
Q9_1	0.69	0.56	0.78
Q9_2	0.90	0.85	0.93
Q9_3	0.95	0.92	0.96
Q9_4	0.60	0.44	0.72
Q10_1	0.51	0.34	0.65
Q10_2	0.61	0.46	0.73
Q12_1	0.44	0.25	0.59
Q12_2	0.53	0.36	0.66
Q12_3	0.39	0.20	0.56
Q12_4	0.43	0.24	0.59
Q12_5	0.50	0.32	0.64
Q12_6	0.52	0.35	0.66
Q13_1	0.45	0.26	0.60
Q13_2	0.63	0.48	0.74
Q13_3	0.44	0.25	0.59
Q13_4	0.58	0.42	0.70
Q14_1	0.63	0.49	0.74
Q14_2	0.54	0.38	0.68
Q14_3	0.52	0.35	0.66
Q14_4	0.50	0.33	0.65
Q14_5	0.56	0.40	0.69
Q29_1	0.64	0.50	0.75
Q29_2	0.68	0.55	0.78
Q29_3	0.63	0.49	0.75
Q16_1	0.61	0.46	0.73
Q16_2	0.60	0.45	0.72
Q16_3	0.42	0.23	0.58
Q16_4	0.61	0.46	0.73
Q16_5	0.78	0.68	0.85
Q16_6	0.41	0.23	0.57
Q16_7	0.43	0.24	0.59
Q16_8	0.30	0.10	0.48
Q16_9	0.48	0.30	0.63
Q16_10	0.25	0.04	0.43
SEE score	0.69	0.56	0.79
SEA score	0.88	0.82	0.92
OE score	0.53	0.36	0.67
IN score	0.57	0.41	0.70
SP score	0.72	0.60	0.81
CA score	0.47	0.29	0.62

See Table 1 for item wording. Lower bound and upper bound reflect the 95 percent confidence interval estimate for the ICC. Score values were obtained by averaging the factor-associated items for each of the 62 students. Abbreviations: *ICC(3,1)* = type 3 intraclass correlation coefficient (two-way, mixed model, single measures), SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

Supplemental Figures

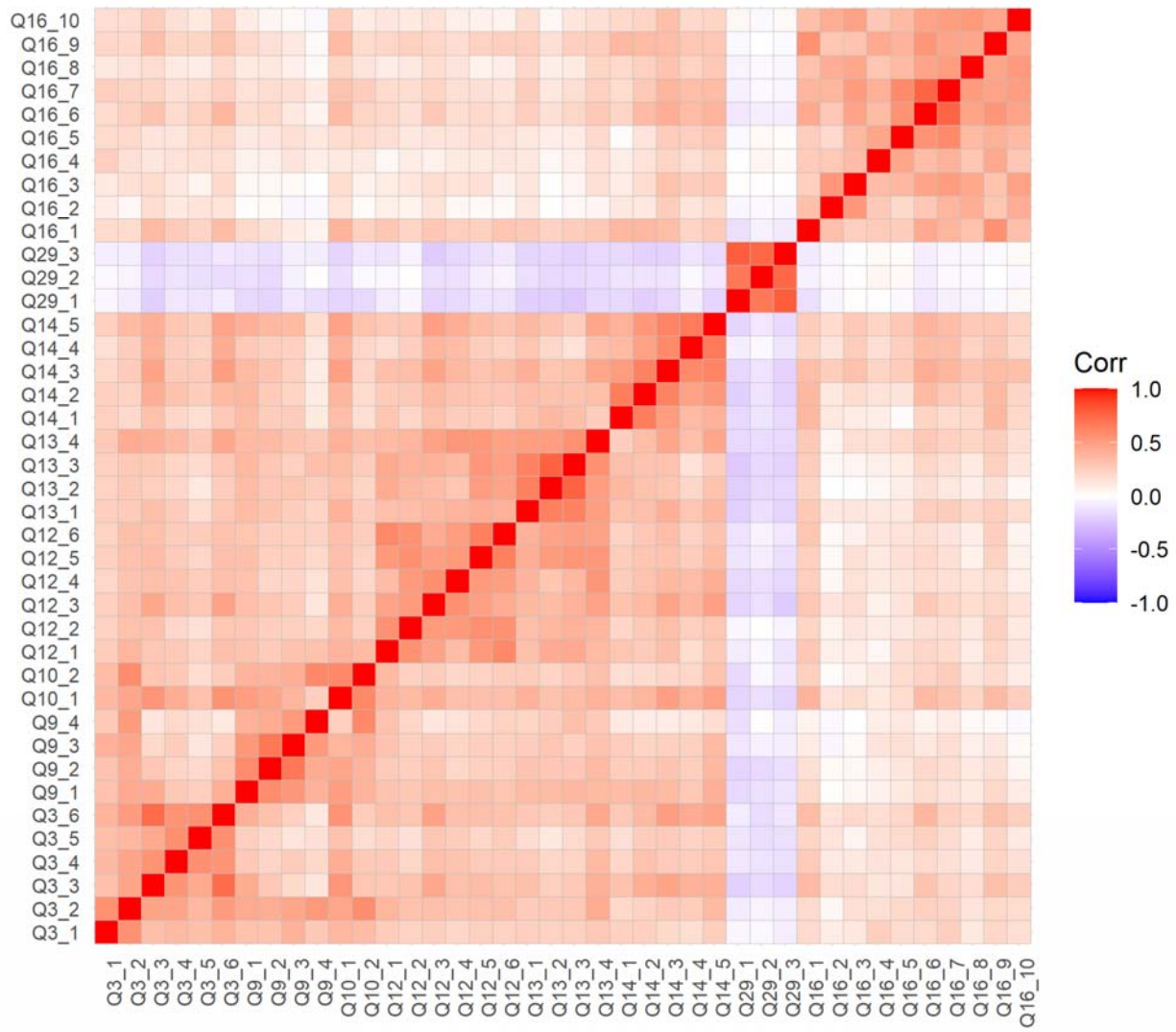


Figure S1. EFA correlation plot

An inter-item correlation plot was used to support factorability of the dataset. Items with positive correlations are shaded in red, while those with negative correlations are blue. Intensity of the color indicates the strength of the correlation. Items with the strongest relationships are clustered together. At least six factors can be identified. Several items exhibit correlations between 0.2 to 0.8. For item wording, see Table 1. *n* = 400.

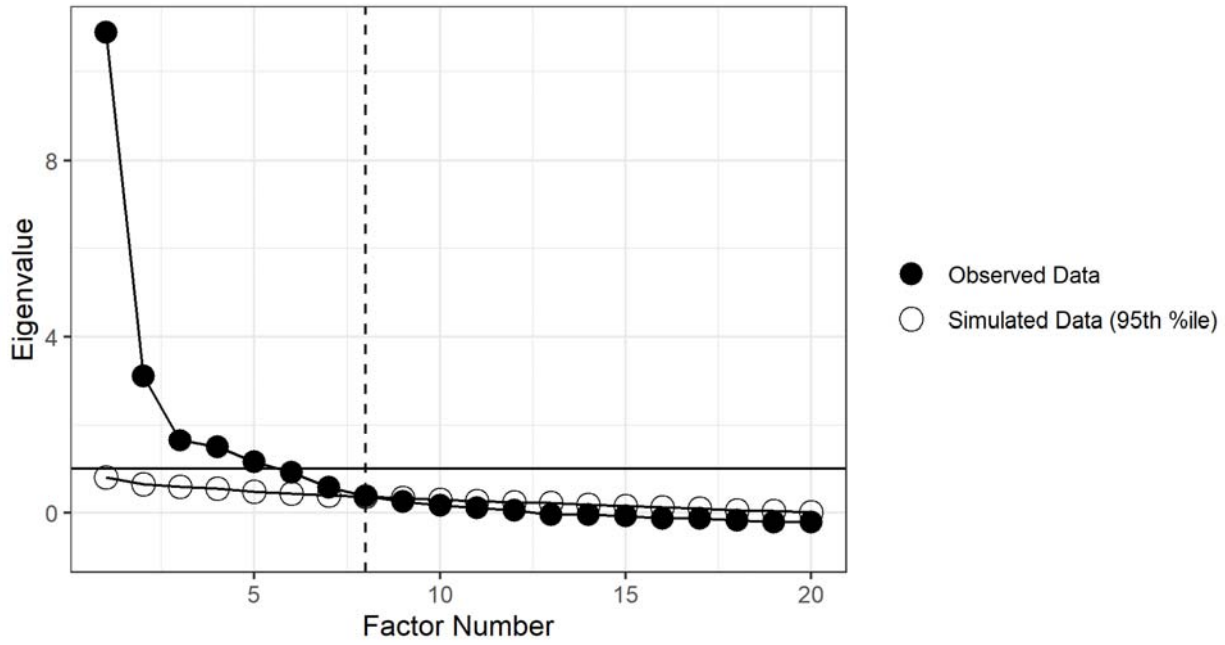
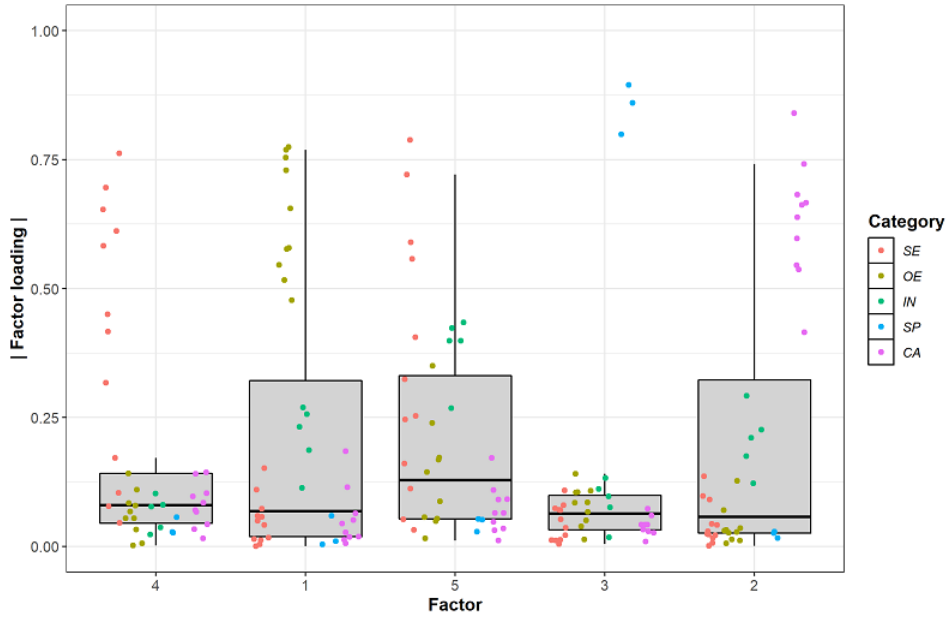


Figure S2. Scree plot

Scree plot of eigenvalues (variance explained; descending order) for each factor that could be extracted from the data (filled circles). Eigenvalues for factors extracted from parallel analysis are also shown (open circles).

A.



B.

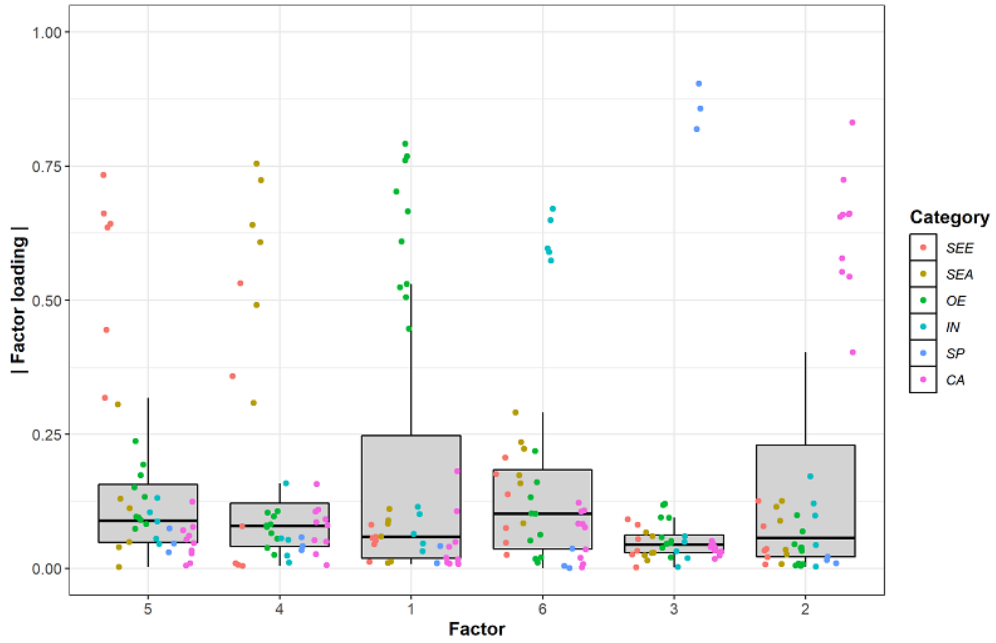


Figure S3. EFA box plots

EFA results were visualized using box plots for the A) five factor and B) six factor models. The absolute value of each item's factor loading was plotted for each factor. Items were colored according to their a priori determined factor. Abbreviations: SE = self-efficacy, SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

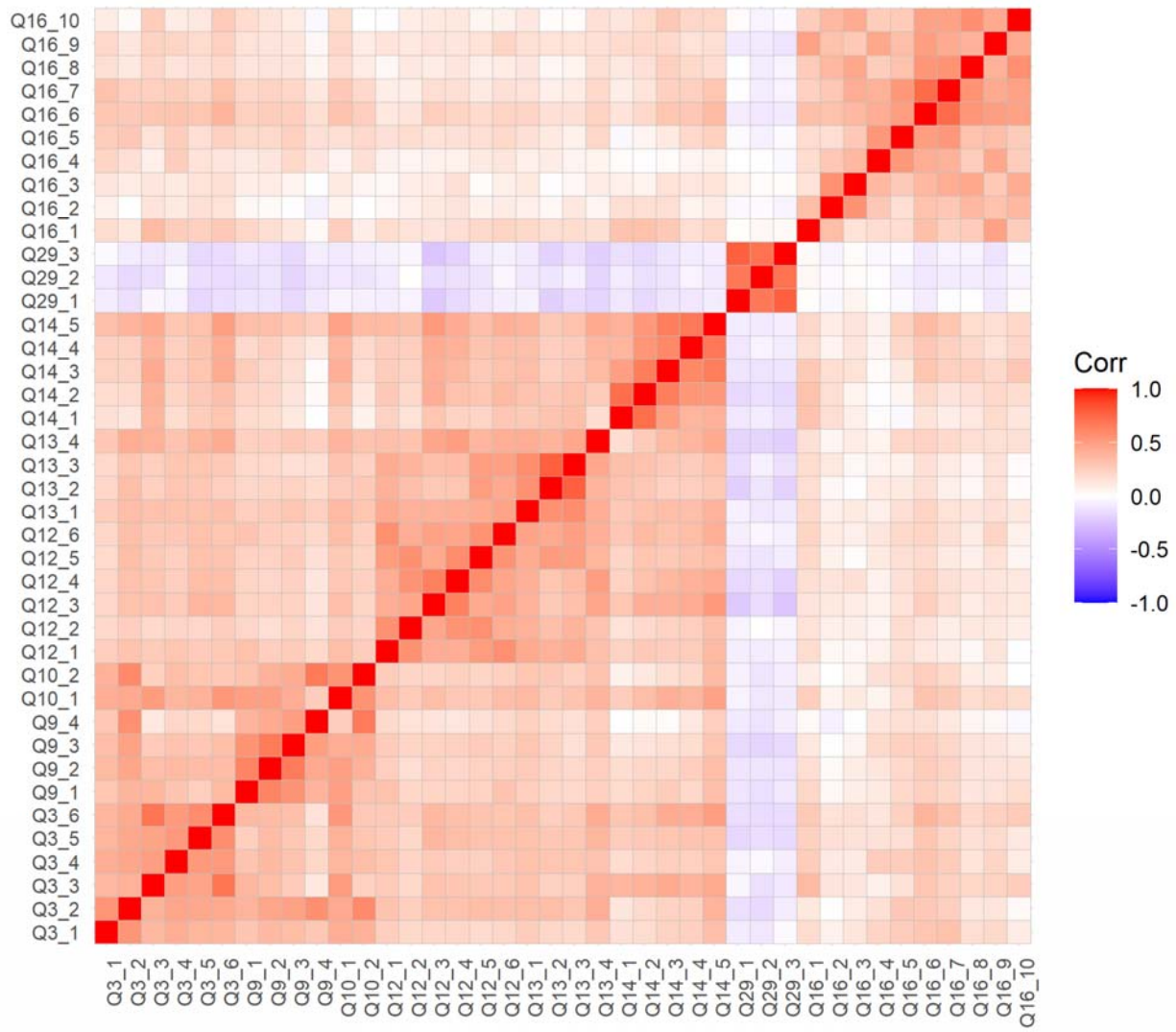


Figure S4. CFA correlation plot

An inter-item correlation plot was used to support factorability of the dataset. Items with positive correlations are shaded in red, while those with negative correlations are blue. Intensity of the color indicates the strength of the correlation. Items with the strongest relationships are clustered together. At least six factors can be identified. Several items exhibit correlations between 0.2 to 0.8. See Table 1 for item wording. *n* = 532.

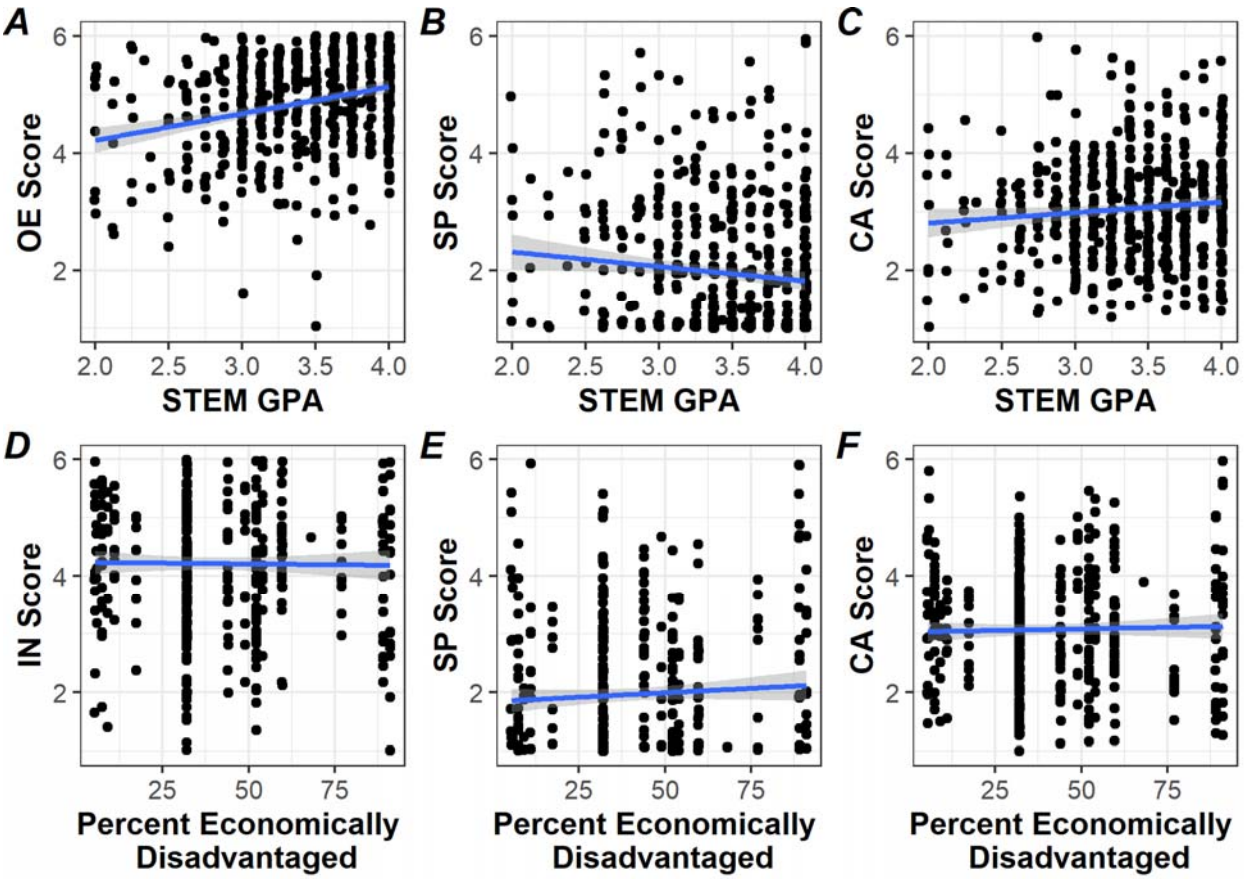


Figure S5. Additional criterion-related validity testing

Scatter plots illustrating the relationship of factor scores and external variables are shown with linear regression lines. Spearman's correlations and *p* values were also calculated for SASS factors with STEM GPA: A) OE score (*rho* = 0.24, $p < 10^{-7}$); B) SP score (*rho* = -0.10, *p* = 0.03); C) CA score (*rho* = 0.09, *p* = 0.03); and for SASS factors with school percent economically disadvantaged: D) IN score (*rho* = -0.01, *p* = 0.8); E) SP score (*rho* = 0.06, *p* = 0.2); and F) CA score (*rho* = 0.02, *p* = 0.6). A high percentage of economically disadvantaged students corresponds with low SES for the school. Because the SES data is collected at the school level, the values are clustered at discrete values on the horizontal axis. Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

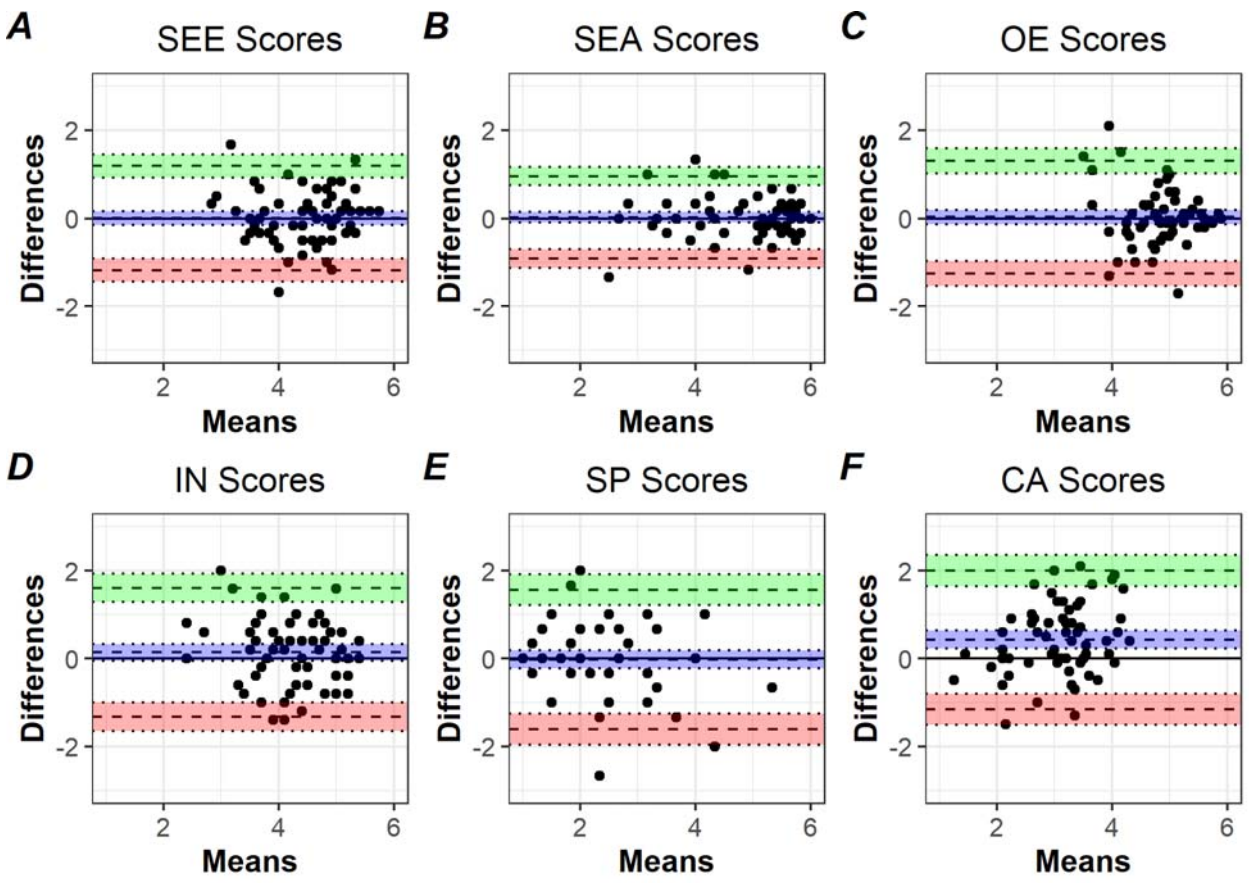


Figure S6. Bland-Altman plots

Bland-Altman plots were generated to evaluate the agreement between student factor scores at each timepoint. For each factor, the differences between each score at each timepoint was calculated and plotted against the mean for each timepoint. The biases, or the average of the distances between the measurements (A: 0.008, B: 0.02, C: 0.03, D: 0.14, E: -0.02, and F: 0.42) and the lower and upper limits of agreement (A: -1.17 and 1.19, B: -0.91 and 0.96, C: -1.25 and 1.31, D: -1.34 and 1.61, E: -1.61 and 1.57, and F: -1.15 and 2.00) are represented as dashed lines. The shaded areas represent 95% confidence intervals for the biases (purple), upper limits of agreement (green), and lower limits of agreement (red). Abbreviations: SEE = self-efficacy experience, SEA = self-efficacy academic, OE = outcome expectations, IN = interests, SP = perceptions of scientists, and CA = career awareness.

References

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, *2*(1), 010101. https://doi.org/10.1103/PhysRevSTPER.2.010101

Aldridge, V. K., Dovey, T. M., & Wade, A. (2017). Assessing test-retest reliability of psychological measures. *European Psychologist*, *22*(4), 207–218. https://doi.org/10.1027/1016-9040/a000298

Anderson, E., & Kim, D. (2006). *Increasing the Success of Minority Students in Science and Technology*. American Council on Education. https://vtechworks.lib.vt.edu/bitstream/handle/10919/86900/MinorityStudentsScience.pdf

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology; New York*, *4*(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*(1), 9–44. https://doi.org/10.1177/0149206311410606

Bartlett, M. S., & Fowler, R. H. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, *160*(901), 268–282. https://doi.org/10.1098/rspa.1937.0109

Bland, M. J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Blotnicky, K. A., Franz-Odendaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career

activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, *5*(1), 22. https://doi.org/10.1186/s40594-018-0118-3

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Cameron, C., Lee, H. Y., Anderson, C., Byars-Winston, A., Baldwin, C. D., & Chang, S. (2015). The role of scientific communication skills in trainees' intention to pursue biomedical research careers: A social cognitive analysis. *CBE Life Sciences Education*, *14*(4), ar46. https://doi.org/10.1187/cbe.14-09-0152

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education*, *98*(3), 443–460. https://doi.org/10.1002/sce.21108

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265. https://doi.org/10.1002/sce.3730670213

Chang, M. J., Eagan, M. K., Lin, M. H., & Hurtado, S. (2011). Considering the impact of racial stigmas and science identity: Persistence among biomedical and behavioral science aspirants. *The Journal of Higher Education*, *82*(5), 564–596. https://doi.org/10.1080/00221546.2011.11777218

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields*. National Center for Education Statistics, Institute of Education Sciences, US Department of Education. https://nces.ed.gov/pubs2014/2014001rev.pdf

Christidou, V. (2011). Interest, attitudes and images related to science: Combining students'

voices with the voices of school Science, teachers, and popular science. *International Journal of Environmental & Science Education*, *6*(2), 141–159.

Cohen, C., Patterson, D. G., Kovarik, D. N., & Chowning, J. T. (2013). Fostering STEM career awareness: Emerging opportunities for teachers. *Washington State Kappan*, *6*(2), 12–17. https://nwabr.org/sites/default/files/pagefiles/PDK%20STEM%20Career%20Awareness.pdf

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7). https://doi.org/10.7275/jyj1-4868

Datta, D. (2017). *blandr: a Bland-Altman method comparison package for R*. https://doi.org/10.5281/zenodo.824514

Dikmenli, M. (2010). Undergraduate biology students' representations of science and the scientist. *College Student Journal*, *44*(2), 579–588.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publications Ltd. https://dl.acm.org/doi/book/10.5555/2502692

Finson, K. D. (2002). Drawing a scientist: What we do and do not know after fifty years of drawings. *School Science and Mathematics*, *102*(7), 335–345. https://doi.org/10.1111/j.1949-8594.2002.tb18217.x

Fouad, N. A., Hackett, G., Smith, P. L., Kantamneni, N., Fitzpatrick, M., Haag, S., & Spencer, D. (2010). Barriers and supports for continuing in mathematics and science: Gender and educational level differences. *Journal of Vocational Behavior*, *77*(3), 361–373. https://doi.org/10.1016/j.jvb.2010.06.004

Fouad, N. A., & Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: Moving the needle. *Journal of Career Assessment*, *25*(1), 24–39. https://doi.org/10.1177/1069072716658324

Garriott, P. O., Flores, L. Y., & Martens, M. P. (2013). Predicting the math/science career goals of low-income prospective first-generation college students. *Journal of Counseling*

Psychology, *60*(2), 200–209. https://doi.org/10.1037/a0032074

Gayles, J. G., & Ampaw, F. (2014). The impact of college experiences on degree completion in STEM fields at four-year institutions: Does gender matter? *The Journal of Higher Education*, *85*(4), 439–468. https://doi.org/10.1080/00221546.2014.11777336

Gore, J., Holmes, K., Smith, M., Southgate, E., & Albright, J. (2015). Socioeconomic status and the career aspirations of Australian school students: Testing enduring assumptions. *The Australian Educational Researcher*, *42*(2), 155–177. https://doi.org/10.1007/s13384-015-0172-5

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*(6), 911–922. https://doi.org/10.1016/j.econedurev.2010.06.010

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265–275. https://doi.org/10.1037/0033-2909.103.2.265

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in Science, Technology, Engineering, and Mathematics*. American Association of University Women. https://time.com/wp-content/uploads/2015/05/why-so-few-women-in-science-technology-engineering-and-mathematics.pdf

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Johnson, L. S. (2000). The relevance of school to career: A study in student awareness. *Journal of Career Development*, *26*(4), 263–276. https://doi.org/10.1177/089484530002600403

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). *semTools: Useful tools for structural equation modeling* (Version 0.5-3) [Computer software]. https://CRAN.R-project.org/package=semTools

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The development of the STEM Career Interest Survey (STEM-CIS). *Research in Science Education*, *44*(3), 461–481. https://doi.org/10.1007/s11165-013-9389-3

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE Life Sciences Education*, *18*(1), rm1. https://doi.org/10.1187/cbe.18-04-0064

Koonce, D. A., Zhou, J., Anderson, C. D., Hening, D. A., & Conley, V. M. (2011). What is STEM? *2011 ASEE Annual Conference & Exposition*, 22.1684.1–22.1684.9. https://doi.org/10.18260/1-2--18582

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kovarik, D. N., Patterson, D. G., Cohen, C., Sanders, E. A., Peterson, K. A., Porter, S. G., & Chowning, J. T. (2013). Bioinformatics education in high school: Implications for promoting science, technology, engineering, and mathematics careers. *CBE Life Sciences Education*, *12*(3), 441–459. https://doi.org/10.1187/cbe.12-11-0193

Krajkovich, J. G., & Smith, J. K. (1982). The development of the Image of Science and Scientists Scale. *Journal of Research in Science Teaching*, *19*(1), 39–44. https://doi.org/10.1002/tea.3660190106

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. https://doi.org/10.1006/jvbe.1994.1027

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. https://doi.org/10.1037//0022-0167.47.1.36

Lent, R. W., Brown, S. D., Sheu, H.-B., Schmidt, J., Brenner, B. R., Gloster, C. S., Wilkins, G., Schmidt, L. C., & Lyons, H. (2005). Social cognitive predictors of academic interests and goals in engineering: Utility for women and students at historically black universities. *Journal of Counseling Psychology*, *52*(1), 84–92. https://doi.org/10.1037/0022-0167.52.1.84

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., & Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: A meta-analytic path analysis of the social-cognitive choice model by gender and race/ethnicity. *Journal of Counseling Psychology*, *65*(1), 17–35. https://doi.org/10.1037/cou0000243

Lim, S. Y., & Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics*, *82*(1), 145–164. https://doi.org/10.1007/s10649-012-9414-x

Lopez, F. G., & Lent, R. W. (1992). Sources of mathematics self-efficacy in high school students. *The Career Development Quarterly*, *41*(1), 3–12. https://doi.org/10.1002/j.2161-0045.1992.tb00350.x

Lopez, F. G., Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Role of social-cognitive expectations in high school students' mathematics-related interest and performance. *Journal of Counseling Psychology*, *44*(1), 44–52. https://doi.org/10.1037/0022-0167.44.1.44

Marshall, C. E., Blalock, C. L., Liu, Y., Pruski, L. A., & Lichtenstein, M. J. (2010). Psychometric re□evaluation of the Image of Science and Scientists Scale (ISSS). *School Science and Mathematics*, *107*(4), 149–160. https://doi.org/10.1111/j.1949-8594.2007.tb17929.x

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing

Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Mau, W.-C., Chen, S.-J., & Lin, C.-C. (2019). Assessing high school student's STEM career interests using a social cognitive framework. *Education Sciences*, *9*(2), 151. https://doi.org/10.3390/educsci9020151

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

Mead, M., & Métraux, R. (1957). Image of the scientist among high-school students: A pilot study. *Science*, *126*(3270), 384–390. https://doi.org/10.1126/science.126.3270.384

Navarro, R. L., Flores, L. Y., & Worthington, R. L. (2007). Mexican American middle school students' goal intentions in mathematics and science: A test of social cognitive career theory. *Journal of Counseling Psychology*, *54*(3), 320–335. https://doi.org/10.1037/0022-0167.54.3.320

Oh, Y. J., Jia, Y., Lorentson, M., & LaBanca, F. (2013). Development of the Educational and Career Interest Scale in science, technology, and mathematics for high school students. *Journal of Science Education and Technology*, *22*(5), 780–790. https://doi.org/10.1007/s10956-012-9430-8

PCAST. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Executive Office of the President. https://files.eric.ed.gov/fulltext/ED541511.pdf

Peterman, K., Withy, K., & Boulay, R. (2018). Validating common measures of self-efficacy and career attitudes within informal health education for middle and high school students. *CBE Life Sciences Education*, *17*(2), ar26. https://doi.org/10.1187/cbe.17-07-0122

Peters, G.-J. Y. (2014). The alpha and the omega of scale reliability and validity. *The European Health Psychologist*, *16*(2), 56–69.

Pion, G. M., & Lipsey, M. W. (1981). Public attitudes toward science and technology: What have

the surveys told us? *Public Opinion Quarterly*, *45*(3), 303–316. https://doi.org/10.1086/268666

Portney, L., & Watkins, M. (2015). *Foundations of clinical research: Applications to practice*. F.A. Davis Company.

Rasheed Ali, S., & Saunders, J. L. (2009). The career aspirations of rural Appalachian high school students. *Journal of Career Assessment*, *17*(2), 172–188. https://doi.org/10.1177/1069072708328897

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria. http://cran.univ-paris1.fr/web/packages/dplR/vignettes/intro-dplR.pdf

Revelle, W. R. (2020). *psych: Procedures for personality and psychological research* (Version 2.0.9) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Roller, S., Lampley, S., Dillihunt, M., Benfield, M., & Turner, M. (2020). Student attitudes toward STEM: A revised instrument of social cognitive career theory constructs (fundamental). *2018 ASEE Annual Conference & Exposition Proceedings*. 2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah. https://doi.org/10.18260/1-2--31001

Rosseel, Y. (2012). *lavaan: An R package for structural equation modeling*. Journal of Statistical Software. https://doi.org/10.18637/jss.v048.i02

Rouquette, A., & Falissard, B. (2011). Sample size requirements for the internal validation of psychiatric scales. *International Journal of Methods in Psychiatric Research*, *20*(4), 235–249. https://doi.org/10.1002/mpr.352

Salonen, A., Kärkkäinen, S., & Keinonen, T. (2018). Career-related instruction promoting students' career awareness and interest towards science learning. *Chemistry Education Research and Practice*, *19*(2), 474–483. https://doi.org/10.1039/C7RP00221A

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger (Ed.), *The role of interest in learning and*

development (Vol. 461, pp. 183–212). Lawrence Erlbaum Associates, Inc, xiv. https://psycnet.apa.org/fulltext/1992-97926-008.pdf

Schinske, J., Cardenas, M., & Kaliangara, J. (2015). Uncovering scientist stereotypes and their relationships with student race and student success in a diverse, community college setting. *CBE Life Sciences Education*, *14*(3), 14:ar35. https://doi.org/10.1187/cbe.14-12-0231

Schneider, L., Jacque, B. H., & Henry, J. (2018). *BioSCANN: A collaborative learning platform that scaffolds scientific inquiry in the context of interrupted case studies*. https://repository.isls.org/bitstream/1/694/1/357.pdf

Schneider, L., Jacque, B., Slotta, J., McDonough, E. K., & Henry, J. (2019, June). *Using digital interrupted case studies for whole class inquiry in life sciences*. https://doi.org/10.22318/cscl2019.636

Scholes, L., & Stahl, G. (2020). "I"m good at science but I don't want to be a scientist': Australian primary school student stereotypes of science and scientists. *International Journal of Inclusive Education*, 1–16. https://doi.org/10.1080/13603116.2020.1751316

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics*, *79*(2), 215–237. https://doi.org/10.1007/s10649-011-9341-2

Semsar, K., Knight, J. K., Birol, G., & Smith, M. K. (2011). The Colorado Learning Attitudes about Science Survey (CLASS) for use in biology. *CBE Life Sciences Education*, *10*(3), 268–278. https://doi.org/10.1187/cbe.10-10-0133

Shaw, E. J., & Barbuti, S. (2010). Patterns of persistence in intended college major with a focus on STEM majors. *NACADA Journal*, *30*(2), 19–34. https://doi.org/10.12930/0271-9517-30.2.19

Sheu, H.-B., & Bordon, J. J. (2017). SCCT research in the international context: Empirical

evidence, future directions, and practical implications. *Journal of Career Assessment*, *25*(1), 58–74. https://doi.org/10.1177/1069072716657826

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Stake, J. E. (2006). The critical mediating role of social encouragement for science motivation and confidence among high school girls and Boys. *Journal of Applied Social Psychology*, *36*(4), 1017–1045. https://doi.org/10.1111/j.0021-9029.2006.00053.x

Syed, M., Goza, B. K., Chemers, M. M., & Zurbriggen, E. L. (2012). Individual differences in preferences for matched-ethnic mentors among high-achieving ethnically diverse adolescents in STEM. *Child Development*, *83*(3), 896–910. https://doi.org/10.1111/j.1467-8624.2012.01744.x

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *The Journal of the International Association of Medical Science Educators: JIAMSE*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

Turner, S. L., Joeng, J. R., Sims, M. D., Dade, S. N., & Reid, M. F. (2019). SES, gender, and STEM career interests, goals, and actions: A test of SCCT. *Journal of Career Assessment*, *27*(1), 134–150. https://doi.org/10.1177/1069072717748665

UMass Donahue Institute. (2011). *Increasing student interest in Science, Technology, Engineering, and Math (STEM): Massachusetts STEM pipeline fund programs using promising practices*. Massachusetts Department of Higher Education. https://www.mass.edu/stem/documents/Student%20Interest%20Summary%20Report.pdf

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.

https://doi.org/10.1177/0734282915571160

van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education*, *41*(2), 150–164. https://doi.org/10.1080/09500693.2018.1540897

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081–1121. https://doi.org/10.3102/0002831213488622

Watson, J. C. (2017). Establishing evidence for internal structure using exploratory factor analysis. *Measurement and Evaluation in Counseling and Development*, *50*(4), 232–238. https://doi.org/10.1080/07481756.2017.1336931

Wei, T., & Simko, V. (2017). *R package "corrplot": Visualization of a correlation matrix* (Version 0.84) [Computer software]. https://github.com/taiyum/corrplot

Westbrook, B. W., Elrod, T., & Wynne, C. (1996). Career assessment and the Cognitive Vocational Maturity Test. *Journal of Career Assessment*, *4*(2), 139–170. https://doi.org/10.1177/106907279600400203

Wyer, M. (2003). Intending to stay: Images of scientists, attitudes toward women, and gender as influences on persistence among science and engineering majors. *Journal of Women and Minorities in Science and Engineering*, *9*(1), 1–16. https://doi.org/10.1615/JWomenMinorScienEng.v9.i1.10

Wyer, M., Schneider, J., Nassar-McMillan, S., & Oliver-Hoyo, M. (2010). Capturing stereotypes: Developing a scale to explore U.S. college students' images of science and scientists. *International Journal of Gender, Science and Technology*, *2*(3). http://genderandset.open.ac.uk/index.php/genderandset/article/view/78/178

Zhang, L., & Barnett, M. (2015). How high school students envision their STEM career pathways. *Cultural Studies of Science Education*, *10*(3), 637–656. https://doi.org/10.1007/s11422-013-9557-9