Validation of a Measure of STEM Interest for Adolescents

Nancy L. Staus [1] · Kristin Lesseig [2] · Richard Lamb [3] · John Falk [4] ·
Lynn Dierking [1]

Abstract

Students' declining performance in science and mathematics is an issue of great international concern. Recently, educators and researchers have begun to focus on affective factors such as interest to better understand STEM learning and persistence. Therefore, there is a need for effective measures of STEM interest that allow it to be tracked over time and to provide opportunities for early interventions by educators. One such instrument was recently developed to measure youth interest in STEM as a general construct and in four domains associated with STEM: earth and space science, life science, technology and engineering, and mathematics. In this paper, we explore the psychometric properties of the measure scales by administering the instrument to a large sample of youth from both a traditional and STEM-focused school and examining theorized relationships using confirmatory factor analysis as part of a larger structural equation model. Results provided support for a single latent dimension of STEM interest, confirmed the existence of the four individual STEM interest dimensions, and provided evidence of structural and generalizability validity. We conclude that the instrument provides a sufficient means to measure STEM interest for adolescent youth within a variety of populations and educational contexts.

Keywords Confirmatory factor analysis · STEM education · STEM interest · Structural equation modeling · Survey instrument

Advancing STEM education to meet current and future social and economic challenges is an urgent goal for many education systems worldwide (English, 2016; Fitzallen, 2015). In particular, US students' performance in science and mathematics compared with youth in other industrialized countries is becoming a matter of great concern and study (National Research Council [NRC], 2011). Several international assessments (e.g. Programme for International Student Assessment (PISA), National Assessment of Educational Progress

✉ Nancy L. Staus
  stausn@oregonstate.edu

Extended author information available on the last page of the article

(NAEP), and Trends in International Mathematics and Science Study (TIMSS)) indicate that students in the USA show declining performance in mathematics and science during the middle school years such that by age 15, US students perform significantly worse in math and science assessments than the youth in most other countries tested (Ahmed, van der Werf, Kuyper, & Minnaert, 2013; National Science Board, 2010). Ultimately, the decline in performance manifests as reductions in the selection of science, technology, engineering, and mathematics (STEM) discipline courses in secondary and post-secondary education (Lamb et al. 2015; Sadler , Sonnert, Hazari, & Tai, 2012).

In order to better understand how to strengthen students' skills in STEM, educators and researchers worldwide are increasingly attending to the affective and socioemotional factors related to STEM content learning including attitudes, interest, and motivation, due to their importance in the learning process (Fortus, 2014; Maltese, Melki, & Wiebke, 2014; Maltese and Tai 2011; Vedder-Weiss & Fortus 2011). In particular, a growing amount of evidence suggests that youth interest in STEM strongly influences persistence over time, even more so than achievement measured with grades and standardized test scores (Maltese et al., 2014; Maltese & Tai 2011; Tai,, Liu, Maltese, & Fan, 2006). Therefore, interest in STEM content and activities during adolescence may be a significant predictor of future engagement in STEM activities or careers.

Consequently, there is a growing focus on identifying strategies for measuring, capturing, and sustaining young people's interest in STEM-related content and activities. For example, a recent report in the USA published by the President's Council of Advisors on Science and Technology called on STEM-related educational fields to "create STEM-related experiences that excite and interest students of all backgrounds" (President's Council of Advisors on Science and Technology [PCAST], 2010, p. v) and there are similar calls for strengthening STEM education worldwide (e.g. Office of the Chief Scientist [OCS], 2013; The Royal Society Science Policy Centre, 2014). But in order to realize the goal of increasing youth interest in STEM, there is a need for appropriate measures of STEM interest that allow it to be tracked over time and to provide opportunities, both in and out of school, for early interventions by educators to address patterns of declining interest.

As part of a longitudinal study of STEM interest pathways in middle school youth, the authors constructed the [Synergies] survey (hereafter, Survey) instrument designed to measure STEM interest both as a general construct and to capture specific interest in the four content domains: science, technology, engineering, and mathematics (Falk et al., 2016). The instrument was administered to 257 youth (ages 10–14) in a single urban middle school and subjected to an exploratory psychometric analysis which identified four underlying factors within the STEM items which were classified as earth and space science, life science, technology and engineering, and mathematics.

The purpose of the current study is to further explore the validity of the survey as an instrument capable of measuring the construct of STEM and the four underlying domains. We examined the psychometric properties of our measure scales by administering the affective measure to a large heterogeneous sample of youth and conducting a confirmatory factor analysis as a part of a larger structural equation model to verify theorized relationships between items and latent constructs that emerged from the exploratory factor analysis. In addition, we administered a number of science- and mathematics-specific interest items based on the ASPIRE survey (DeWitt et al., 2011) simultaneously with the STEM items to investigate

generalizability validity of the instrument. Our results indicated that the survey can be used to meaningfully compare youth interest in STEM and four STEM domains between and across groups of youth.

Theoretical Framework

A problematic issue for researchers and educators is the lack of an agreed-upon definition for STEM education, as it has been interpreted in a variety of ways (e.g. Burke, Francis, & Shanahan, 2014; English 2016; Moore and Smith, 2014). Our definition of STEM was informed by that of the US National Science Foundation (NSF) who originally coined the term to account for the fact that problems are often not easily divisible into separate disciplines such as physics or biochemistry (NRC, 2012), suggesting that STEM itself can be thought of as a meta-discipline that is a new "whole" (Morrison, 2006). However, in practice, most schools in the USA and elsewhere continue to teach STEM disciplines individually (e.g. in separate science, math, and technology classes). As new education standards such as the Next Generation Science Standards (NGSS) in the USA and others worldwide (e.g. Australian Curriculum, Assessment and Reporting Authority [ACARA], 2015; UK Department for Education 2015) call for the teaching of cross-cutting concepts in addition to standard science and engineering content (NRC 2013), it will become increasingly necessary to have tools to measure interest in the four domains of STEM as well as STEM as an integrated discipline.

Our conceptualization of STEM interest was guided by the work of Hidi and Renninger (2006) and Renninger and Su (2012) in which interest development is conceptualized as a multiple-phase process during which early "situational interests" may become well-developed "individual interests" over time through repeated engagement with the topic or activity and with support and encouragement from others. Situational interest is an early phase of interest that is externally triggered, consists mostly of affect, and may be fleeting in nature. Consequently, early situational interest may be difficult if not impossible to measure because the learner may not even be aware of the interest (Renninger & Hidi 2011). However, if the early situational interest persists, it may develop over time into an individual interest, which is a relatively enduring and self-motivated predisposition to engage with and learn about a specific topic or activity. Therefore, it is likely that it is these more well-developed interests that are captured through a survey instrument.

Survey development was also influenced by the person-object theory of interest (Krapp, 2002), which posits that interest is always specific to a certain content or activity and develops through the interaction between a person and their environment. In other words, one can only become interested in topics or activities that one is aware of and has opportunities in which to engage. This suggests that when measuring interest, it is important to know what opportunities are likely to be available to the individuals in the study population. In particular, in order to assess STEM interests that develop in adolescents through their interactions with STEM learning resources and opportunities in their community both in and out of school, the survey instrument needed to include a variety of STEM-related content items that were accessible to youth in the study population.

Finally, STEM interest appears to be a topic rather than domain specific (Krapp, 2002; Krapp & Prenzel, 2011). In other words, youth may have a strong interest in specific subject areas such as electricity while reporting very low interest in the corresponding science domain (i.e. physics). In addition, STEM interest is often gendered based on the domain (Krapp & Prenzel, 2011; Sjøberg & Schreiner, 2010). For example, Häussler and Hoffmann (2002) found that while interest in biology or the life sciences was just as pronounced in girls as in boys, if not more so, girls showed less interest than boys in physics and chemistry subject areas. Similarly, in their multi-country study of science education, Sjøberg & Schreiner (2010) found that girls' and boys' science and technology interests were highly context-dependent with boys reporting more interest in topics that were technical, mechanical, electrical, violent, and explosive, while girls were most interested in topics associated with health, medicine, and the human body. These findings suggest that the terminology used when constructing survey items matters a great deal, and using broad domain terms such as "science" or "technology" may not fully capture youth interest in underlying topics that they may not associate with those broader domains.

Survey Instrument Development

The purpose of this paper is to further explore the psychometric properties and potential general usefulness of an existing survey instrument designed to examine youth interest in STEM. A complete discussion of survey development and exploratory psychometric analysis is available elsewhere (Falk et al., 2016). Here we provide a summary of the survey development process to allow the reader to better understand how and why survey items were chosen and included in the final instrument.

Item Development

The survey instrument was originally developed for use in the longitudinal [Synergies] project in which we examined the STEM interest pathways of youth from ages 10/11 to 13/14 in an urban community located in a metropolitan area in the Northwestern USA (Falk et al., 2016). Development of the instrument was motivated in part by the lack of existing instruments that measured outcomes for multiple STEM disciplines simultaneously. A recent review of STEM-focused assessments (Minner, Erickson, Wu, & Martinez, 2012) found that of the 69% of instruments that measured the cognitive dimensions of STEM, 53% focused exclusively on science. Of the remaining instruments, 26% focused exclusively on mathematics, with only 21% attempting to measure science and mathematics interest in an integrated manner. None of the instruments measuring STEM-related outcomes in the psychosocial domains, such as attitudes, interest, or motivation, addressed all STEM domains concurrently.

Our goal was to develop a self-reporting instrument that measured interest in STEM as an integrated or meta-discipline (e.g. Morrison, 2006) as well as measuring interest in each of the separate disciplines (e.g. science, math, engineering, and technology) that make up the larger construct of STEM. To increase content validity, our construct was theoretically grounded in and informed by the four-phase model of interest (Hidi & Renninger, 2006) and person-object theory of interest (Krapp, 2002) as discussed above and thoroughly reviewed by our project advisors, several of whom were experts

in the fields of science and STEM interest development. Because STEM interest is a topic rather than domain specific (Krapp, 2002; Krapp & Prenzel, 2011), we avoided the use of domain-level terms that might have a negative connotation to some youth (e.g. math), might not be understood (e.g. engineering), or may be so general that they were likely to be interpreted differently by different youth (e.g. science). Instead, STEM was operationally defined as a curated assemblage of youth-focused activities and/or practices (i.e. taking things apart, solving puzzles, etc.) related to STEM that the youth in this community would recognize and have opportunities in which to engage. In addition, we included a variety of items that are known to appeal to both girls (e.g. life science items) and boys (e.g. technology items) to address the gendered nature of STEM interest (Sjøberg & Schreiner 2010).

Because of the importance of the environment to interest development (Krapp, 2002), we purposely narrowed our content focus to those areas that our target cohort of youth was most likely to encounter in school (e.g. components of life and earth sciences) and out of school (e.g. consumer technology and gardening). While several items were adapted from the Relevance of Science Education Questionnaire (ROSE; Schreiner & Sjøberg, 2004), most were developed to capture the STEM opportunities available to youth in [Sunrise Middle School[1]]. Eight adolescent youth from the community were hired as researchers to help develop and review survey items to ensure they represented topics and activities that were available to most youth. For example, rather than asking the youth if they were interested in "technology," we asked about specific practices that were technology-related and part of the everyday experience of this group of youth such as interest in "how computers or cell phones work"

Table 1 Summary of STEM interest items and associated domains that emerged from the exploratory PCA

STEM interest items	STEM domain
What it is like on other planets and exploring space	Earth and space science
How stars and planets form	Earth and space science
Why clouds, rain, and weather happen	Earth and space science
How earthquakes, volcanoes, and hurricanes happen	Earth and space science
What to eat and how to exercise to keep healthy and fit	Life science
How traits are passed from parents to children	Life science
How the human body works	Life science
How buildings and bridges are made	Technology and engineering
How computers and cell phones work	Technology and engineering
How to use and make maps	Technology and engineering
How to design new games or toys	Technology and engineering
How gas and diesel engines work	Technology and engineering
How to do Sudoku or other math problems	Mathematics
How to measure the size or area of things	Mathematics
How to solve puzzles	Mathematics
How to make different shapes and patterns out of things	Mathematics

Items coded on a five-point scale from 1 = "Dislike a lot" to 5 = "Like a lot" in response to the prompt: "How much do you like finding out about"

(for the list of final items see Table 1). By specifically identifying areas already familiar to the participants, survey items allowed the youth to more accurately identify with STEM applications and content. In addition, we chose science survey items that specifically related to the science curriculum of [Sunrise Middle School] middle school attended by the majority of youth in the [Synergies] project. This process culminated in the development of 23 items representing a diversity of STEM content and practices that youth might encounter in their daily lives. While it encompassed a broad range of STEM topics, it was not an exhaustive list as we did not include STEM domains that youth had few opportunities with which to engage (e.g. physics). In addition, we were constrained by the survey length and associated completion time limitations required by such instruments.

Exploratory Analysis

We used the principal components analysis (PCA) and a measure of internal consistency (Cronbach's alpha) to investigate the psychometric properties of our instrument. Items with factor loadings under 0.5 or that lowered the reliability of the scale as measured by Cronbach's alpha were removed (Beavers et al., 2013). The removal of items resulted in a 16-item STEM interest measure, three to five interest items per content construct (Table 1). The four identifiable STEM interest components which emerged from the PCA were earth and space science, life science, technology and engineering, and mathematics. After confirming the internal consistency reliability of each content interest scale using Cronbach's alpha, the authors computed a composite reliability score for the overall measure of STEM interest (Fornell & Larcker, 1981; Raykov, 1998). These scores represented each youth's STEM interest as a point measure.

Methods

Survey Instrument

As described above, the survey instrument used in this study consisted of 16 STEM interest items representing four STEM domains that emerged from the exploratory PCA. In addition to these items, we included 17 science interest items from the ASPIRE questionnaire (DeWitt et al., 2011) and 17 mathematics interest items created by substituting the word "math" for "science" in the ASPIRE items (Table 2). Overall, these items were internally consistent and reliable (Rasch's reliability = 0.80 and 0.84 for science and math items, respectively). By including interest measures focusing exclusively on mathematics and science, we were able to examine the relationship between STEM interest, science interest, and mathematics interest in order to help establish instrument validity through convergent validity analysis (Messick, 1989).

Participants

In order to provide evidence of validity and reliability, we obtained data from 811 youth, grades 6–8, in two schools in a metropolitan area in the Northwestern USA. One school was a traditional public middle school while the other was a STEM-designated

Table 2 Summary of science/mathematics interest items and associated domains adapted from the ASPIRE survey (DeWitt et al. 2011)

Science/math interest items[1]
I enjoy learning science/math
I am pretty good at science/math
I learn things quickly in science/math
I like science/math
I find science/math to be really interesting
Science/math is boring[2]
I find science/math difficult
Science/math is harder for me than for other kids my age
I am just not good at science/math
My parents are interested in science/math
My parents want me to be interested in science/math
My parents would be happy if I became a scientist/mathematician someday
Science/math will be useful in my future
I see how science/math relates to my life
Science/math is helpful in understanding today's world
My science/math teacher makes me interested
Science/math is useful in helping solve problems

[1] Items coded on a five-point scale from 1 = "Strongly disagree" to 5 = "Strongly agree"

[2] Item reverse coded for analysis

school. The two participating schools were chosen because of connections with mathematics and science education researchers at a local university. [Ridgeview STEM Academy] is an inclusive, STEM-focused school in the largest district in the region. Students submit an application for entry and a lottery system is used to obtain a student population that mirrors district-wide achievement test scores and demographics. Project-based learning (Buck Institute, n.d.) is a key component of the school's vision, and students all have access to technology as a design, research, and communication tool. Teachers collaborate throughout the year to develop interdisciplinary projects, using overarching themes to integrate the humanities and STEM disciplines (Lesseig et al., 2019). Demographic data for both schools are presented in Table 3. At the time of the STEM interest survey, the school was just beginning its third year. Enrollment was approximately 60 students per grade level.

In contrast, [Hood Middle School] is a traditional middle school in the second largest district in the same region. A small group of teachers expressed interest in interdisciplinary projects and were currently in year one of a three-year STEM professional development project. However, as is typical of US middle schools, the school still adhered to typical middle school structures wherein the content was taught in isolated courses and cross-disciplinary teacher collaboration was limited.

Study participant responses were collected for validation and item reduction purposes over the course of the school year. Ninety-four percent of all students in the sample completed the measure sufficiently for psychometric analysis purposes.

Table 3 Demographics of youth in participating schools

Demographic data	[Ridgeview STEM Academy]	[Hood Middle School]
Total student enrollment	170	875
Minority population	26% (12% Hispanic, 6% 2 or more races, 3% Asian, 3% Black)	44% (28% Hispanic, 6% 2 or more races, 3% Asian, 4% Black)
Gender	Male 68% Female 32%	Male 53% Female 47%
Free or reduced meal eligibility	29%	55%
8th grade state test proficiency	71% science 45% mathematics	58% science 29% mathematics

Forty-eight participants were removed from the analysis due to their failure to complete all survey questions resulting in a final N of 763. Analysis of non-responding participants suggested that they did not differ significantly in characteristics or traits when compared with the whole sample. Participants who completed the survey were 49% female and 51% male (Table 3). The grade break was approximately equal across the sample with one-third of the sample coming from each grade (sixth, seventh, and eighth). A comparison between the demographic characteristics and current US census data seems to indicate an over-representation of groups (e.g. Hispanics). However, sample analysis using a hypergeometric distribution test does not result in a statistically significant overrepresentation by any particular group $p = 0.074$ (Harkness, 1965).

Analysis

The authors developed a structural equation model consisting of a path analysis and a confirmatory factor analysis using Mplus 7.11 to determine the relationships and predictability between the constructs of mathematics interest, science interest, and STEM discipline content interest constructs. The Mplus code is available from the authors upon request. Each of the models separately and in combination not only provides evidence of the relationship within each measure but also between the measured constructs via a path analysis which indicates the role and place of science and mathematics interest related to STEM discipline interest content outcomes. A Multiple Indicator Multiple Cause Structural Equation Modeling (MIMIC) approach was used for the development of a proposed model of the interaction between mathematics interest, science interest, and STEM discipline content interest. The MIMIC model is a general approach to multivariate data analysis in which the purpose is to study the complex, causal relationships among unobserved variables as a system. The primary advantages of SEM over traditional multiple regression are (a) more flexibility in the assumptions, (b) the use of confirmatory factor analysis to reduce measurement error, (c) the test of multiple dependent variables, (d) measuring direct and indirect effects including error, and (e) providing superior results with missing, time-series, autocorrelated, and non-normal data (Ruo et al., 2008).

Results

Structural Equation Model

The structural equation model analysis of the interaction between mathematics interest, science interest, and STEM discipline content interest illustrates that the data has an excellent model fit, $X2(46) = 204.45$, $p < 0.001$, RMSEA = 0.046 CI 90% = [0.001, 0.008], CFI = 0.983, TLI = 0.977 (Hu & Bentler, 1999). The authors calculated measure reliability using the latent trait reliability method (LTRM) in addition to Cronbach's alpha because this method does not have the limitations associated with Cronbach's alpha (Dimitrov, 2012). The measurement reliability is calculated at 0.85 using LTRM and 0.87 using a composite Cronbach's alpha. These levels of internal consistency are considered sufficient for a measure of this type (Raykov, 1998). The resulting MIMIC model is over-identified suggesting there are other rival models that can be examined. However, the examination of a rival model fit the data shown in Table 4 via log-likelihood; AIC and BIC suggest that the proposed model is the most parsimonious of the models.

Item Parceling

Item parcels differ from subscales or scale scores in that the entire set of item parcels reflects a single latent construct (Cattell, 1956). Item parcels are preferred in this analysis as indicators as they are often more reliable indicators and more normally distributed. Scores of the item parcels are also often more continuous in nature; in addition, they require a smaller sample and avoid the problem of less than three indicators per construct in the confirmatory component of the SEM. The use of item parceling also increases the resolution of second-order constructs within a SEM (Marsh, Morin, Parker, & Kaur, 2014). The resulting confirmatory model establishes the presence of three latent traits, mathematics interest, science interest, and STEM discipline content, significantly covarying ($\xi = 0.628$) with each other and causal paths via mathematics and science (λ Mathematics = 0.69, λ Science = 0.72).

The authors did not include the residual variance of error for the latent variable and items for clarity purposes. The residual variance of errors for each item and variable range from $\Theta = 0.09$ to $\Theta = 0.32$. Tests for measurement and structural invariance across the model suggest that the measured items meet the assumptions of invariance of factor loadings. Figure 1 illustrates the confirmatory model of the relationship between interest items, their constructs, and STEM disciplines. Collectively, mathematics and science interests are most predictive of the ESS (Earth and Space Science) subscale, followed by the TE (technology and engineering) subscale.

Table 4 MIMIC SEM model comparisons

Estimate	Model 1 (proposed)	Model 2 (rival 1)	Model 3 (rival 2)
Log-likelihood	− 322.489	− 242.47	− 131.27
AIC	1.298	0.986	0.826
BIC	− 2449.90	− 2597.51	− 3211.01

The structural coefficient with the path from "math to science" (0.62) is statistically significant thus indicating that science interest moderates mathematics interest on STEM content interest outcomes. Without the interest in science, mathematics interest reduces its impact on STEM interest on average by 4%. Due to the significant intercorrelation coefficient between each of the STEM discipline subscales, a second order latent trait (integrated STEM) is justified. After the addition of STEM to the model, no significant change in model fit was observed ($\Delta X2(4) = 0.24$, $p = 0.12$).

Discussion

The purpose of this study was to further examine the relationship between the constructs and properties of a freestanding, self-reporting instrument designed to measure youth interest in STEM as a general construct and interest in four domains associated
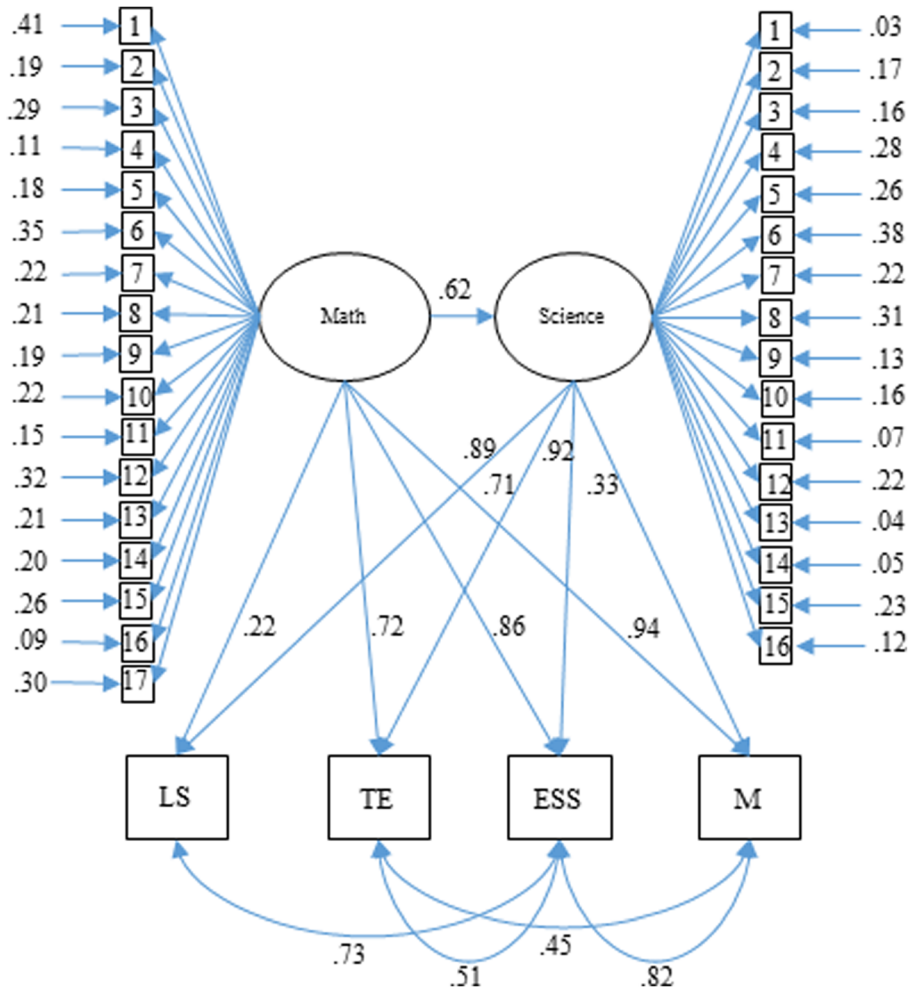

Fig. 1 Structural equation model of mathematics interest, science interest, STEM disciplines

with STEM: earth and space science, life science, technology and engineering, and mathematics. The results provided strong support for a single latent dimension of STEM interest underlying the responses. In addition, the confirmatory factor analysis (CFA) confirmed the existence of the four individual STEM interest dimensions. Although there were multiple models available for consideration, the model presented here was the most parsimonious and consistent with the findings of the exploratory analysis (Falk et al., 2016).

The validity of the STEM components was further supported by the relationships between the science and mathematics interest scales and the four STEM interest domains. For example, results indicated that general mathematics interest was most predictive of interest in the math STEM domain ($r = 0.94$) and least predictive of life science interest ($r = 0.22$) as illustrated in the path diagram (Fig. 1). In addition, general science interest was most strongly predictive of life science ($r = 0.89$) and earth and space science interest ($r = 0.92$), and least related to the math STEM domain ($r = 0.33$). These relationships provided evidence for the generalizability aspect of Messick's framework of validity (1989) by comparing the instruments that measure similar STEM interest domains. Generalizability and convergent validity were reinforced by the fact that the exploratory and confirmatory analyses produced the same results in terms of constructs and item relationships. Evidence of the validity of the four STEM domains is also illustrated via the fact that the analyses were performed on data from two very different samples of youth. The participants in the [Synergies] project were from an urban area and were largely low-income and ethnically diverse, whereas those at [Hood Middle School] and [Ridgeview STEM Academy] were from a suburban population and half were Caucasian. In addition, the inclusion of both a traditional and STEM-focused school in the analysis provides evidence of validity for different educational approaches. Together, these findings indicate that the instrument has the potential to be a useful measure of STEM interest in a variety of communities and contexts.

Another noteworthy finding was that science interest acted as a moderator of mathematics interest as it related to STEM interest as illustrated in the path diagram (Fig. 1). In other words, the relationship between mathematics interest and STEM interest increased as science interest increased. These findings are particularly relevant in light of the growing call for integrative efforts in STEM education in which two or more STEM subjects are taught simultaneously so that STEM learning becomes more connected and meaningful for learners (Becker & Park, 2011; NRC, 2012; Sanders, 2009). A recent meta-analysis investigating the effects of integrative approaches to STEM education revealed that broadly speaking, such approaches appear to have positive effects on student learning outcomes (Becker & Park, 2011), particularly for mathematics education. For example, Judson and Sawada (2000) found significantly higher math achievement for students in an integrated science and math class. Others found more positive attitudes toward mathematics in interdisciplinary courses (Elliot, Oly, McArthur, & Clark, 2001). Our findings provide additional support for integrating science and math instruction which may increase youth interest in STEM.

The relationships among individual STEM dimensions were less clear-cut. While the strong correlation ($r = 0.73$) between life science and earth and space science was expected, the stronger correlation ($r = 0.82$) between math and earth and space science is more difficult to explain. In addition, we would have expected a stronger relationship between mathematics and technology and engineering. This could be due to the fact that the topics/

activities we chose to include in the survey instrument largely focused on using consumer technology (e.g. how cell phones work) rather than technology development (e.g. how to code video games) which would require more mathematical interest and capabilities. Thus, research instruments highlighting different facets of technology and engineering might yield different results. However, it is also likely that the varying strengths of these relationships were driven by the fact that most adolescent youth do not fully understand the connections among STEM domains in real-world applications. Again, these findings point to the potential efficacy of integrated, multi-disciplinary classes and activities that explicitly link related STEM dimensions such as mathematics and engineering. For example, one study at the college level showed that integrating engineering and mathematics led to greater learning outcomes as well as a better understanding of why they needed to know both content areas (Everett, Imbrie, & Morgan, 2000).

Conclusions

The results of this study indicate that the STEM interest survey is a valid and reliable measure of interest in the construct known as STEM as well as interest in four distinct STEM domains: earth and space science, life science, technology and engineering, and mathematics. In addition, the robustness of our model indicates that interest scores will not fluctuate due to random variation, and the results of differential item functioning indicate invariance across the sample. Therefore, we conclude that the instrument provides a sufficient means to measure STEM interest that will allow educators and researchers to measure interest in four STEM domains with a high level of reliability and validity across a variety of populations and learning contexts. By measuring changes in these interest scores over time, educators will be able to develop appropriate interventions to address declining interest or differences in STEM interest for different groups of youth (e.g. by gender, ethnicity). Such customized interventions may help educators to better support STEM interest development during adolescence for more youth.

References

Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology, 105*(1), 150–161.

Australian Curriculum, Assessment and Reporting Authority. (2015). *The Australian Curriculum, Version 7.3, January 5, 2015*. Sydney, NSW: ACARA.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation, 18*(6), 1–13.

Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education, 12*, 23–37.

Buck Institute. (n.d.). *What is project-based learning (PBL)?* Retrieved 24 October 2018, from http://www.bie. org/about/what_pbl.

Burke, L., Francis, K., & Shanahan, M. (2014, July). *A horizon of possibilities: A definition of STEM education*. Paper presented at the STEM 2014 Conference, Vancouver, Canada.

Cattell, R. B. (1956). Personality and motivation theory based on structural measurement. In J. L. McCary (Ed.), *Psychology of personality* (pp. 51–98). New York, NY: Logos.

DeWitt, J., Osborne, J., Archer, L., Dillon, J., Willis, B., & Wong, B. (2011). Young children's aspiration in science: The unequivocal, the uncertain and the unthinkable. *International Journal of Science Education, 35*(6), 1037–1063.

Dimitrov, D. (2012) *Statistical methods for validation of assessment scale data in counseling and related fields.* Alexandria, VA: American Counseling Association.

Elliot, B., Oly, K., McArthur, J., & Clark, B. (2001). The effect of an interdisciplinary algebra/science course on students' problem solving skills, critical thinking skills and attitudes towards mathematics. *International Journal of Mathematical Education in Science and Technology, 32*(6), 811–816.

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

Falk, J. H., Staus, N. L., Dierking, L. D., Penuel, W., Wyld, J., & Bailey, D. (2016). Understanding youth STEM interest pathways within a single community: the Synergies project. *International Journal of Science Education, Part B, 6*(4), 369-384.

Everett, L. J., Imbrie, P. K., & Morgan, J. (2000). Integrated curricula: Purpose and design. *Journal of Engineering Education, 89*(2), 167–175.

Fitzallen, N. (2015). STEM education: What does mathematics have to offer? In M. Marshman (Eds.), *Mathematics education in the margins. Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia* (pp. 237–244). Sydney, Australia: MERGA.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Fortus, D. (2014). Attending to affect. *Journal of Research in Science Teaching, 51*(7), 821–835.

Harkness, W. (1965). Properties of extended hypergeometric distributions. *The Annals of Mathematical Statistics, 36*(3), 938–945.

Häussler, P., & Hoffmann, L. (2002). An intervention study to enhance girls' interest, self- concept, and achievement in physics classes. *Journal of Research in Science Teaching, 39*(9), 870–888.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Judson, E., & Sawada, D. (2000). Examining the effects of a reformed junior high school science class on students' math achievement. *School Science and Mathematics, 100*(8), 419–425.

Krapp, A. (2002). An educational-psychological theory of interest and its relation to self- determination theory. In E. Deci & R. Ryan (Eds.), *The handbook of self-determination research* (pp. 405–427). Rochester, NY: University of Rochester Press.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education, 33*(1), 27–50.

Lamb, R., Akmal, T., & Petrie, K. (2015). Development of a cognition-priming model describing learning in a STEM classroom. Journal of Research in Science Teaching, 52(3), 410-437.

Lesseig, K., Firestone, J., Morrison, J., Slavit, D., & Nelson, T. (2019). An analysis of cultural influences on STEM schools: Similarities and differences across K-12 contexts. *International Journal of Science and Mathematics Education. 17*(3), 449-466.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education, 95*(5), 877–907.

Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education, 98*, 937–962.

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: Macmillan.

Minner, D., Erickson, E., Wu, S., & Martinez, A. (2012). *Compendium of research instruments for STEM education. Part 2: Measuring students' content knowledge, reasoning skills, and psychological attributes.* Cambridge, MA: ABT Associates. Retrieved on July 27, 2014 from http://cadrek12.org/announcements/findingsdr-k-12-stem-school-study-s3-project Retrieved July 27, 2014.

Moore, T. J., & Smith, K. A. (2014). Advancing the state of the art of STEM integration. *Journal of STEM Education, 15*(1), 5–10.

Morrison, J. (2006). Attributes of STEM education: *The students, the academy, the classroom*. Baltimore, MD: Teaching Institute for Excellence in STEM.

National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

National Research Council. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.

National Science Board. (2010). *Science and engineering indicators 2010*. Arlington, VA: National Science Foundation.

Office of the Chief Scientist. (2013). *Science, technology, engineering and mathematics in the national interest: A strategic approach*. Canberra, Australia: Australian Government. Canberra: Australian Government.

President's Council of Advisors on Science and Technology. (2010). *Prepare and inspire: K-12 education in science, technology, engineering and math (STEM) for America's future*. Washington, DC: Executive Office of the President.

Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement, 22*(4), 375–385.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184.

Renninger, K. A., & Su, S. (2012). Interest and its development. In R. Ryan (Ed.), The Oxford handbook of human motivation (pp. 167–187). New York, NY: Oxford University Press.

Ruo, B., Baker, D. W., Thompson, J. A., Murray, P. K., Huber, G. M., & Sudano, J. J., Jr. (2008). Patients with worse mental health report more physical limitations after adjustment for physical performance. *Psychosomatic Medicine, 70*(4), 417–421.

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education, 96*(3), 411–427.

Sanders, M. (2009). STEM, STEM education, STEM mania. *Technology Teacher, 68*(4), 20–26.

Schreiner, C., & Sjøberg, S. (2004). Sowing the seeds of ROSE. Background, rationale, questionnaire development and data collection for ROSE (The Relevance of Science Education) – A comparative study of students' views of science and science education (pdf) (Acta Didactica 4/2004). Oslo, Norway: Department of Teacher Education and School Development, University of Oslo.

Sjøberg, S., & Schreiner, C. (2010). The ROSE project. An overview and key findings. Retrieved from. http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf.

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*, 1143–1144.

The Royal Society Science Policy Centre. (2014). *Vision for science and mathematics education*. London, England: The Royal Society.

UK Department for Education (2015). National curriculum in England: Science programmes of study. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study.

Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not? *Journal of Research in Science Teaching, 48*(2), 199–216.

Affiliations

Nancy L. Staus[1] · **Kristin Lesseig**[2] · **Richard Lamb**[3] · **John Falk**[4] · **Lynn Dierking**[1]

Kristin Lesseig
kristin.lesseig@vancouver.wsu.edu

Richard Lamb
rllamb@buffalo.edu

John Falk
john.falk@freechoicelearning.org

Lynn Dierking
dierkingl@science.oregonstate.edu

1   Oregon State University, Corvallis, OR, USA

2   Washington State University, Vancouver, WA, USA

3   University at Buffalo, Buffalo, NY, USA

4   Institute for Learning Innovation, Portland, OR, USA