IJEMST

International Journal of Education in Mathematics, Science and Technology

ISSN: 2147-611X

International Journal of Education in Mathematics, Science and Technology (IJEMST)

www.ijemst.com

Instruments to Measure Elementary Student Mindsets about Smartness and Failure in General and with respect to Engineering

Pamela S. Lottero-Perdue[1], **Cathy P. Lachapelle**[2]
[1]Towson University
[2]Museum of Science, Boston

To cite this article:

Lottero-Perdue, P.S. & Lachapelle, C. P. (2019). Instruments to measure elementary student mindsets about smartness and failure in general and with respect to engineering. *International Journal of Education in Mathematics, Science and Technology (IJEMST), 7*(2), 197-214. DOI:10.18404/ijemst.552468

International Journal of Education in Mathematics, Science and Technology

Volume 7, Number 2, 2019 DOI:10.18404/ijemst.552468

Instruments to Measure Elementary Student Mindsets about Smartness and Failure in General and with respect to Engineering

Pamela S. Lottero-Perdue, Cathy P. Lachapelle

Article Info

Article History

Received:
21 January 2019

Accepted:
21 March 2019

Keywords

Elementary school
Engineering education
Smartness
Failure
Mindset

Abstract

The aim of this study was to assess evidence for the validity of General Mindset (GM) and Engineering Mindset (EM) surveys that we developed for fifth-grade students (ages 10-11). In both surveys, we used six items to measure student mindset to determine if it was more fixed (presuming intelligence is fixed and failure is a sign that one is not smart enough) or more growth-minded (presuming one can become smarter and that failures are signals to improve) (Dweck, 1986). We administered surveys to 2473 fifth-grade students (ages 10-11) who learned one or two engineering units during one academic year. Using Exploratory Factor Analysis (EFA) then Confirmatory Factor Analysis (CFA), we identified a single factor for each survey. We assert that there is strong evidence for the validity of using the GM or EM survey with students ages 10-11. The EM survey should be given after students have engaged in engineering classwork in school.

Introduction

The purpose of this study is to assess evidence for the validity of two surveys we developed to measure the general and engineering mindsets, respectively, of fifth-grade students (ages 10-11). In what follows, we first articulate the particular type of mindset to which we refer, why it is important to measure in elementary students, and how it is measured. We then connect this to our context of focus: engineering education.

Growth and Fixed Mindset

According to a theory first proposed by Dweck (1986, 1999, 2006), students have one of two basic mindsets about themselves as learners when they approach a learning task or experience: 1) a fixed mindset, also referred to as an entity theory of intelligence; or 2) a growth mindset, also referred to as an incremental theory of intelligence. Those who have a fixed mindset presume that they have "a predetermined amount of intelligence, skills or talents" in general or in a particular area (Ricci, 2013, p. 3). Experiencing struggle, difficulty, or failures suggests to those with a fixed mindset that they are simply not smart, skilled, or talented enough. Those with a growth mindset believe that they can become smarter, more skilled, or more talented with effort, persistence and new approaches to problems (Dweck, 2015). Growth-minded action involves: learning from the struggles, difficulties or failures; trying again (and again) with a new and better approach to the task; and, ultimately, improving. Students may be inclined towards a growth or fixed mindset in general and may have a different mindset in particular domains (e.g., reading, mathematics). There are multiple and complex reasons that students may have domain-specific growth mindsets (Matheson, 2015; Stipek & Gralinski, 1996), including the influence of cultural norms, expectations and stereotypes. For example, those who are underrepresented within science, mathematics or engineering by gender may receive gender-essentialist fixed mindset messages such as boys are good at science (Wonch Hill et al., 2017) or girls are not good at math (Tomasetto, Alparone, & Cadinu, 2011).

Elementary Student Mindset

Elementary student mindset is of importance because multiple studies have found that, as a group, elementary students with a growth mindset have more positive academic and attitudinal outcomes than students with a fixed mindset (Gunderson, Sorhagen, Gripshover, Dweck, Goldin-Meadow, and Levine, 2013; McCutchen, Jones, Carbonneau & Mueller, 2016; Park, Gunderson, Tsukayama, Levine & Beilock, 2016; Petscher, Al Otaiba, Wanzek, Rivas and Jones, 2017). These findings are consistent with seminal mindset studies of students in middle and high-school and at the university level (e.g., Blackwell, Trzesniewski, and Dweck, 2007; Good, Aronson & Inzlicht, 2003; Good, Rattan & Dweck, 2012; Yeager et al., 2016). Additionally, there is evidence from this group of older students – middle school and beyond – that instructional interventions may support the development of growth mindset in students (e.g., Blackwell, Trzesniewski, and Dweck, 2007; Good, Aronson & Inzlicht, 2003; Good, Rattan & Dweck, 2012; Paunesku et al., 2015, Yeager et al., 2016).

Thus far, there is little clarity in the literature about how general and domain-specific mindsets differ, no matter the age of the student. A common finding across three studies in the elementary literature was that there is no difference (McCutchen et al., 2016; Park et al., 2016; Stipek & Gralinski, 1996). However, Petcher and colleagues (2017) found that the relationship between mindset and reading achievement was strongest when general and reading-specific mindsets were combined in the model to predict reading achievement.

Measuring Elementary Student Mindset

General and domain-specific mindset is commonly measured through the use of surveys. Elementary mindset surveys utilize items that originate from either: 1) a survey by Dweck (1999); or 2) a survey by Gunderson and colleagues (2013) with origins in Heyman and Dweck (1998).

Dweck's 1999 Survey and Related Mindset Surveys

Dweck's (1999) theory of intelligence survey for children aged 10 and older contained six items. These items were evaluated on a six-point Likert scale: 1 = strongly agree, 2= agree, 3 = mostly agree, 4 = mostly disagree, 5 = disagree, 6 = strongly disagree. The first three statements were written from a fixed perspective so that stronger agreement corresponds to a stronger fixed mindset. The latter three are written from a growth perspective. The statements were:

1. You have a certain amount of intelligence, and you really can't do much to change it.
2. Your intelligence is something about you that you can't change very much.
3. You can learn new things, but you can't really change your basic intelligence.
4. No matter who you are, you can change your intelligence a lot.
5. You can always greatly change how intelligent you are.
6. No matter how much intelligence you have, you can always change it quite a bit. (p. Dweck, 1999, p. 177).

Dweck reduced these intelligence items to four, excluding items 1 and 4, in a 2006 publication, and included a separate survey to assess mindset based on "personal qualities" (e.g., "you are a certain kind of person, and there is not much that can be done to really change that") (p. 13). The first three of the Dweck (1999) items were used with 12 other items (e.g., self-efficacy) in an instrument validation study by Hanson (2017) with third through eighth graders. Factor analysis suggested four factors, one of which was what Hanson called the "individual mindset scale" comprised of those three mindset items. Chronbach's α, for the subscale for these three fixed-perspective items was 0.77. These items are also the sole items in an online mindset survey by the Project for Education Research that Scales (PERTS, 2015).

Petscher and colleagues (2017) included the six items from Dweck (1999) in their initial list of 26 items on their mindset survey given to fourth-grade students, but only retained one of them (#3, above) in their final 15-item survey after conducting factor analysis. McCutchen and colleagues (2016) reported using a slight modification of Dweck's (1999) fixed perspective items in their study of fourth through sixth graders, with three statements written within the domain of math (α = .65) and three in reading (α = .71). For example, two items were: "Your smarts in math/reading is something about which you can't change very much" (p. 209).

Mindset Surveys related to Gunderson and Colleagues' (2013) Work

Gunderson et al. (2013) took a somewhat different approach, following the work of Heyman and Dweck (1998) with second and third-grade children, rephrasing items with respect to others (not "you") and replacing words like "intelligence" with "smart." For example, one item was: "Imagine a kid who thinks that a person is a certain amount smart and stays pretty much the same. How much do you agree with this kid?" Gunderson and colleagues used a five-point scale, ranging from 1 = a little to 5 = a lot. The scale had a pictorial element in which children pointed to dots that ranged from the smallest (a little) to the largest dot (a lot). Their survey had a total of 24 items, including items about general and domain-specific mindset. Relevant to the present study, they identified a 14-item subscore that "assessed the belief that traits are fixed versus malleable" ($\alpha = .65$) (2013, p. 400).

Others have followed Gunderson and colleagues' approach with regard to item writing (e.g., "Imagine a kid …") and a 5-point dot-based Likert scale. Schroder and colleagues' (2017) study of gaming by first and second-grade students included eight items written from both fixed and growth perspectives about general mindset (α = 0.72). Park and colleagues' (2016) study of first graders included three fixed-perspective statements about smartness both in general and with respect to mathematics and spelling. Park et al. used McDonald's omega to assess reliability and found it adequate ($\omega = 0.70$ in the fall and 0.82 in the spring of the same academic year) (2016).

Engineering Education and Failure

The domain of focus for this study is engineering. Engineering is a regular feature of elementary science education in the United States, in large part due to the *Next Generation Science Standards* (NGSS Lead States, 2013), which include science and engineering standards and practices. Elementary students learn to engineer by using an engineering design process to solve problems (Cunningham, 2018). There are multiple engineering design processes that can be used. Common elements of design processes include that students: define or are given a problem, constraints and criteria; brainstorm multiple possible solutions to the problem; choose an idea and plan it (aka "the design); physically create and test the design; analyze the design performance against criteria; troubleshoot or improve their design based on this analysis; and iterate, repeating the design process (Crismond & Adams, 2012; Moore et al., 2014). Elementary versions of design processes contain these elements, but simplify them into a short number of steps, e.g.: Ask, Imagine, Plan, Create, Improve (EiE, 2019).

The iterative, improvement-focused nature of the engineering design process suggests its potential alignment with growth mindset (Lottero-Perdue & Parry, 2017a, 2017b). When designs are tested and fail to meet criteria, the desired response in engineering is a growth-minded one: figure out what went wrong, persist, try again in a newer and smarter way (Lottero-Perdue, 2015; Lottero-Perdue & Parry, 2017a, 2017b). Fixed-minded responses like giving up are not productive.

Purpose

As a part of a larger project that encompasses this study, we were motivated to explore the extent to which students who had learned to engineer had a growth or a fixed mindset about engineering; we also wanted to know how demographic or instructional variables may impact students' engineering mindset. However, no instruments had been developed to measure engineering mindset. Beyond this, we wanted to be able to compare engineering to general mindset and to determine how general mindset may predict engineering outcomes. Thus, we created a similarly-worded pair of surveys: a General Mindset (GM) survey and an Engineering Mindset (EM) survey. Our research questions were as follows:

1. What is the evidence for the validity of using the GM survey for measuring elementary students' general mindsets?
2. What is the evidence for the validity of using the EM survey for measuring elementary students' engineering mindsets?

Method

Research Context

The GM and EM surveys were developed as part of the larger Exploring the Efficacy of Engineering is Elementary (E4) Project, a two-year experimental, randomized control study that examines the impact of a treatment and comparison curriculum on 14,015 third, fourth, and fifth-grade students' interests, attitudes, and learning related to engineering and science.

Instrument Development

We began exploratory instrument development work during the first year of data collection for the E4 Project in five third and fourth-grade classrooms with 112 students. This was a convenience sample associated with the classrooms that the first author was observing. We created a single survey that used four general mindset and four engineering mindset items. The general mindset items were from Dweck (2006). Recall that Dweck (2006) included two versions of a four-item mindset survey: 1) about intelligence, and 2) about personal qualities. We chose the personal qualities set, e.g., "You are a certain kind of person, and there is not much that can be done to change that" or "You can always change basic things about the kind of person you are." Our thinking was that personal qualities language may be more accessible to elementary students than intelligence language. In creating the engineering mindset items, we decided to mirror the personal qualities set replacing phrases about "the kind of person you are" with "how good you are at engineering." For example, two of the questions were: "You can learn new things about engineering, but you can't really change how good you are at engineering" or "You can change how good you are at engineering a lot." Students responded via a six-point Likert scale of agreement as done in Dweck (1999). The survey was given after the students completed engineering instruction.

From this pilot investigation, we learned by answering students' questions and observing their completion of the surveys that personal qualities language was (in some students' words) "weird"; they expressed that they weren't sure how to answer the questions. Many also suggested that the general and engineering items were redundant. The general mindset items were on the front of the survey paper and the engineering mindsets were on the back. When they turned the page over from front to back, some thought that they were being asked the same kinds of questions, despite the shift to engineering domain-specific phrases.

In what follows, we describe what we learned from the pilot study, how we adapted mindset surveys by Dweck (1999, 2006), and how we added our own knowledge of engineering design failure to develop the GM and EM surveys. Each section represents a decision we made with regard to the development of the surveys and the items within them.

Separating GM from EM

One of the decisions we made after the pilot study was that general mindset items would be given on a pre-engineering instruction survey (GM survey) and engineering mindset items would be given after engineering instruction (EM survey). There were two reasons for this. First, we wanted to avoid the situation where students felt that the engineering items were the same as the general mindset items they just answered. Second, we did not want to overload students with too many post-instruction questions, since there were already many post-instruction assessments for the entire E4 Project. Regarding this decision to separate the surveys, although it is not reasonable to ask students about engineering mindset prior to learning about engineering, it *is* possible to ask students about their general mindsets prior to learning about engineering. Importantly, this does not mean that the GM survey was a pre-survey in a repeated measures sense (Ruel, Wagner III & Gillespie, 2016).

Focusing on Smartness in General and for Engineering

Another shift that we made from the pilot year to the present study was to use intelligence instead of personal quality items. We aimed to make intelligence the focus of both general and engineering mindset items. Another shift was to simplify the language, using "smart" in place of "intelligent," similar to other studies that have altered mindset survey items for use in elementary classrooms (Park et al., 2016; Petscher et al., 2017; Stipek & Gralinski, 1996). In an attempt to avoid using the word "smartness" in an item, we used "No matter who you are …" instead of "No matter how much intelligence/smartness you have …" language in items G3 and E3 (Table 1). Both versions of this item were included on the Dweck (1999) survey, while the Dweck (2006) survey only used the intelligence/smartness version. Table 1 shows the smartness items for the GM and EM surveys as compared to the items on surveys by Dweck (1999, 2006) surveys. Note that two of each type of smartness items are written from a fixed perspective (G1, E1, G4, E4) and two are written from a growth perspective (G2, E2, G3, E3).

Table 1. Smartness items in the GM, EM, and Dweck (1999, 2006) surveys

GM Survey Items	EM Survey Items	Dweck Survey Items
G1. You can't really change how smart you are.	E1. You can't really change how smart you are at engineering.	2. Your intelligence is something about you that you can't change very much. (Dweck, 1999, 2006)
G2. No matter how smart you are now, you can always become a lot smarter.	E2. No matter how smart you are now at engineering, you can always become a lot smarter.	5. You can always greatly change how intelligent you are. (Dweck, 1999, 2006)
G3. No matter who you are, you can become a lot smarter.	E3. No matter who you are, you can become a lot smarter at engineering.	4. No matter who you are, you can change your intelligence a lot. (Dweck, 1999)
G4. You can learn new things, but you can't really change how smart you are.	E4. You can learn new things about engineering, but you can't really change how smart you are at engineering.	3. You can learn new things, but you can't really change your basic intelligence. (Dweck, 1999, 2006)

Starting from Dweck's item language, we developed the precise language within items G1 through G4 and E1 through E4 through multiple rounds of feedback with members of the E4 Project team, including its PI. The team has extensive experience writing elementary-level items for STEM assessments (e.g., Lachapelle & Brennan, 2018), working with elementary students, and writing elementary curricula. Additionally, we kept the language patterns of the EM survey items as close as possible to those in the GM survey. This way of constructing engineering items is similar to approaches by others studying domain-specific mindset (McCutchen et al., 2016; Park et al., 2016; Petscher et al., 2017; Stipek & Gralinski, 1996).

Adding Failure Items

A concern of the E4 Project team was that the total number of items on the general and engineering surveys needed to be minimized since the project already contained many pre- and post-instruction surveys and assessments. To address this, we aimed to follow Dweck (2006), keeping smartness items to four. We had room to include two questions on each survey that would situate fail words (e.g., fail, fails) within the items (Table 2).

Table 2. Failure items in the General Mindset (GM) and Engineering Mindset (EM) surveys

GM Survey Items	EM Survey Items
G5. If you try and fail at something, that means you are not smart at that kind of thing.	E5. If your design fails, that means you are not smart at engineering.
G6. If you try and fail at something, you would want to try to do that thing again.	E6. If your design fails, you would want to engineer a new design.

These are totally new items to mindset surveys despite the omnipresence of discussions about failure with regard to mindset (i.e., those with a fixed-minded respond to failures in different ways than do those with growth mindset) (Dweck, 1986, 1999, 2006). Further, and as discussed previously, design failure is a normal and expected part of engaging in an engineering design process. Statements about failure were informed by previous work (Lottero-Perdue, 2015; Lottero-Perdue & Parry, 2017a, 2017b), and after conducting focus group interviews with students, part of which included discussions about design failure experiences. This work contributes important evidence for validity of the items for the use of determining mindset and failure orientation of children in grades 3-5 (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014), as the content behind the items was established through qualitative research.

Four-Point Likert Scale

Although Dweck's (1999) survey contained a six-point Likert scale, we chose to reduce to four options: strongly agree, agree, disagree, strongly disagree. Part of the reasoning for this was to maintain consistency with the other survey instruments that the larger E4 Project had students complete at the same time as the mindset survey; all other E4 Project surveys employed a four-point Likert scale. Additionally, the author and project team wanted to reduce the number of response options to simplify the survey; a common approach when administering surveys to elementary students, albeit one that has received some scrutiny (Adelson & McCoach, 2010). Further, it was not desirable to reduce to a five-point Likert scale, which would force a neutral midpoint; there are no neutral midpoints in any of the previous mindset studies that we reviewed. The four points on the scale are labeled "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree."

Fifth Grade Student Participants

Finally, we chose to give the survey to the oldest students in the E4 Project, fifth graders, to give students the best chance of understanding the meaning of the statements. Given the older age of these students, coupled with the 10 or 20 hours spent learning about engineering prior to answering the EM survey, it is likely that the students understood the meaning of the statements on this survey. See the appendix for a copy of the final GM and EM instruments that we used.

Instrument Validation Methods

In this study, we present evidence for the validity of use of the GM and EM surveys to measure student mindset before and after an engineering curriculum intervention. Evidence includes focus groups and interviews with students in grades three through five, collected during the first year of the E4 Project, and quantitative evidence derived from factor analysis of a sample of surveys collected from fifth-grade students during the second year of the E4 Project.

Aspects of Validity Examined

As mentioned above, during development of the GM and EM surveys, we generated items based upon both a literature review of prior surveys assessing mindset and upon prior qualitative research (i.e., on failure) conducted by the authors. This work provides strong content-oriented evidence that the survey validly represents the constructs of interest and establishes a theoretical basis for the surveys. We discussed the target concepts (mindset and failure) with students in the first year of the EE study to establish that the content was appropriate to this age level and understood by students who have engaged in engineering instruction.

We collected qualitative data from a sample of students in the target demographic (American students in grades three through five) as evidence towards establishing the validity of using the mindset surveys with this population. The pilot study and interviews with students about their responses to it provided evidence that students interpret the items as intended and informed the redesign of the survey questions. For evidence regarding internal structure, we use EFA and CFA to establish that the internal structure of the survey and derived factors are related to latent constructs of interest (mindset and failure). All these forms of evidence are recommended before advocating that a particular survey is valid for use in a specific context (AERA, APA, & NCME, 2014; Douglas & Purzer, 2015).

Participant Demographics—Student Mindset Surveys

We collected mindset survey data from fifth-grade students who participated in the second year of the E4 Project (N = 2473) for reasons shared below. Demographics for the fifth-grade study sample were similar to those for the entire study, described in Lachapelle and Brennan (2018), which includes students from a wide range of ethnic, racial, and socio-economic status (SES) groups, from urban, suburban, and rural areas in three noncontiguous states in the eastern United States (Table 3).

Table 3. Demographics of the study sample

	Variables	Students (N=2473) N	% Sample
Gender	Male	1261	51%
	Female	1195	48%
	Missing	*17*	*1%*
Racial/ethnic representation within Engineering Field (by Race/Ethnicity)	White (represented in engineering)	1396	56%
	Asian (represented)	80	3%
	Black (underrepresented)	377	15%
	Hispanic (underrepresented)	253	10%
	Other (e.g. multiracial, Native American; underrepresented)	188	8%
	Missing	*179*	*7%*
Special Education Services via an IEP	Does not have an IEP	1424	58%
	Has an IEP	220	9%
	Missing	*829*	*34%*
SES Measure 1: Eligibility for Free and Reduced Lunch (FRL)*	Not eligible for FRL	617	25%
	Eligible for FRL	606	25%
	Missing	*442*	*51%*
SES Measure 2: Number of Books in the Home*	Few (0-10 books)	257	10%
	One shelf (11-25 books)	424	17%
	One bookcase (26-100 books)	739	30%
	Several bookcases (>100 books)	764	31%
	Missing	*289*	*12%*

* SES is a latent construct that is best measured by multiple variables, most notably parental income, education, and occupation (Duncan, Featherman, & Duncan, 1974; Sirin, 2005; NCES, 2012). Another measure is home resources available to the student (e.g., number of books in the home) (NCES, 2012; Sirin, 2005).

Participant Mean Responses—Student Mindset Surveys

Mean scores on the GM and EM surveys (unimputed data) are given in Table 4, as context for the analysis to be presented. Item 1 from each survey, "You can't really change how smart you are [at engineering]," has the largest variance. Fixed-mindset items (1, 4, and 5) show least agreement overall (means of 1.31 to 1.82) while growth mindset items (2, 3, and 6) show strongest agreement (means of 3.38 to 3.63). Recall that a score of 1 is equivalent to "Disagree," while a score of four matches the label "Strongly Agree" on the survey. Students in our sample of fifth graders tend towards agreement with growth mindset items.

Table 4. Participant mean responses per item

Item	N	Minimum	Maximum	Mean	Std. Deviation
G1	2041	1	4	1.84	.977
G2	2030	1	4	3.61	.585
G3	2010	1	4	3.51	.657
G4	2022	1	4	1.82	.855
G5	2032	1	4	1.67	.775
G6	2029	1	4	3.38	.771
E1	1964	1	4	1.72	.926
E2	1961	1	4	3.63	.637
E3	1950	1	4	3.49	.724
E4	1948	1	4	1.79	.862
E5	1959	1	4	1.31	.679
E6	1963	1	4	3.38	.814

Exploratory Factor Analysis

We conducted Exploratory Factor Analysis (EFA) on a randomly selected half of the sample using MPlus 8.1 (Muthén & Muthén, 1998-2017). The purpose of EFA is to explore the relationship between: interrelated items, such as those observed using a survey; the error associated with each item; and latent (unobserved) constructs that the observed items can be said to measure in some way. We split the data set into equal random halves to provide a separate, independent sample for conducting Confirmatory Factor Analysis (CFA) to confirm the factor structure identified by EFA (Bandalos & Finney, 2010). Bandalos & Finney (2010) recommend that analysts use multiple methods to estimate the possible number of factors before beginning factor analysis; multiple possible factor structures should be tested and compared, with weight given to factor structures that are theoretically plausible. In examining the content of the surveys, we identified two plausible latent structures: 1) the items for each survey may correspond to a single "Mindset" variable (a 1-factor solution), or 2) the items may correspond to two variables, "Smartness" and "Failure" (a 2-factor solution). The quantitative methods we used to predict the number of factors included parallel analysis, in which sample data eigenvalues are compared to eigenvalues from randomly generated data, and the examination of scree plots. We accomplished both methods using a script from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html (O'Connor, 2000) in SPSS 24 (IBM, 2016).

We specified the robust weighted least squares estimator as it is most appropriate for estimating non-normal categorical data. We expected the sample size of each random half to be sufficient, even if extracted communalities are low, because the ratio of sample size (~1000) to expected factors (<5) is quite high (>200:1; MacCallum, Widaman, Zhang, & Hong, 1999). We used the oblique Geomin rotation, the default for EFA with categorical dependent variables in Mplus since we predicted that resulting factors would be correlated. Next, we examined the model results for item loadings and cross-loadings. An item was considered to be loading on a factor if the 2-tailed p-value for its coefficient was <.05. One concern we had in approaching the analysis of the survey items was the existence of items with parallel wording. Items 1 and 4 in each survey contain the phrase "you can't really change how smart you are," Items 2 and 3 contain the phrase "you can (always) become a lot smarter," and Items 5 and 6 both contain the phrase "if you try and fail at something" (GM survey) or "if your design fails" (EM survey). Test items with similar wording can lead to correlated error, also called correlated uniqueness (CU), which can lead to inflated estimates of covariation (if left unspecified in modeling) and the extraction of factors that do not have a basis in theory (Brown, 2006; Marsh, 1996). Therefore, during the EFA, we tested all three sets of possible CU.

We used multiple goodness-of-fit measures to compare models so that a variety of plausible candidate factor solutions could be compared before proceeding to the CFA. We also used fit statistics from each of the categories of absolute fit and comparative fit measures (Kelloway, 2015). Absolute fit measures test the congruence of the covariance matrices for the model as compared to the baseline data. We examined three absolute fit measures according to standard rules of thumb: 1) standardized root mean square residual (SRMR) should be <.80; 2) the root mean square error of approximation (RMSEA) should be <.05; and 3) the $\chi 2$ statistic should show a difference between the fitted model and the baseline model at p<.05. The comparative fit index (CFI) should be >.95 for a model to be considered a good fit (Kelloway, 2015). Most importantly, we evaluated all candidate models for interpretability.

Confirmatory Factor Analysis

For CFA, the relationships of all items to latent variables must be specified in advance since the purpose of CFA is to provide confirmatory evidence of the validity of the theorized model. Once a model for each of the GM and EM surveys was established using EFA, we ran a CFA model with the EFA-determined structures using the half of the sample not used for EFA. We used the robust weighted least squares estimator as estimator and oblique Geomin rotation for all CFA models, and we used the same goodness-of-fit statistics as with EFA. We checked parameter estimates for significance and interpretability (Brown, 2006).

We inspected standard errors for excessively large values, which would indicate an unreliable parameter estimate. For items that do not cross-load on multiple factors, the completely standardized factor loading represents the correlation between item and factor, and the R^2 represents the proportion of variance of the item that is explained by the factor (the communality). We examined factor loading sizes and R^2 values for further evidence of meaningful item-factor relationships. Once a factor model was confirmed, we used Mplus 8.1 to output factor scores for each student. Mplus uses the regression method (DiStefano, Zhu, & Mindrila, 2009) to predict factor scores for each individual participant.

Item-Response Theory (IRT)

To gather information about the reliability of the factors, we examined the test information functions for the scores. Because the factor scores are derived from categorical indicator variables, the reliability (in terms of precision of measurement) will vary as a function of the score (Edelen & Reeve, 2007). For this reason, among others, Cronbach's alpha is not an appropriate measure of reliability for this application (Sijtsma, 2009). We used Mplus 8.1 to output the test information functions for each factor score and calculated the standard error of measurement (Edelen & Reeve, 2007).

Results and Discussion

Exploratory Factor Analysis

Our first step was to determine the number of factors to explore with EFA. Parallel analysis on the full dataset with a 95% probability cutoff indicated three possible factors for the GM items. However, our examination of the scree plot showed a bend at the second factor, indicating two factors. Results for the EM survey items were similar; the scree plot was nearly identical (Figure 1 and Table 5).

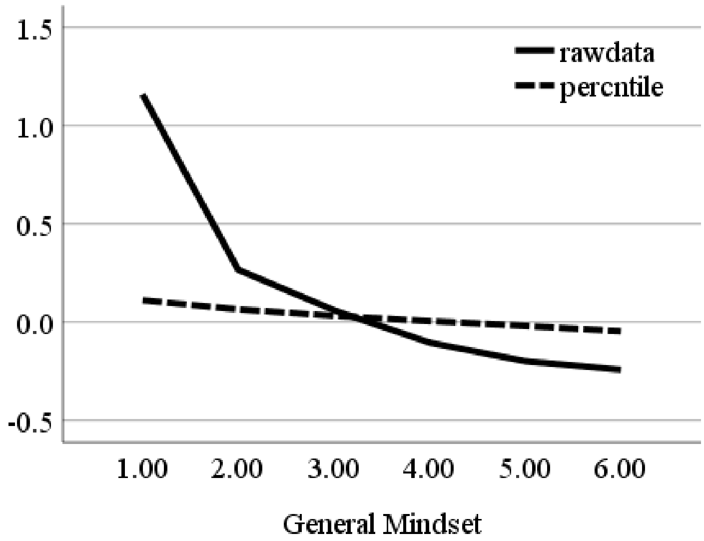

Figure 1. Scree plot

Parallel analysis showed the eigenvalue for the third factor was very similar to that for the randomly generated data. We considered that three factors for six items would greatly reduce the degrees of freedom, negatively impacting the analysis, and that theoretically we had identified only one- and two-factor possible structures. We decided to explore factor structures with one and two factors, forgoing a three-factor analysis.

Table 5. Parallel analysis for GM and EM items

	GM Eigenvalues			EM Eigenvalues		
Factor	Sample Data	Random Data Mean	Random Data Percentile	Sample Data	Random Data Mean	Random Data Percentile
1	1.322162	.079383	.111024	1.159785	.078211	.111135
2	.240423	.042694	.065239	.266294	.041885	.064710
3	.071179	.013837	.031745	.060491	.013340	.032119
4	124598	-.010852	.004767	-.104500	-.010907	.005080
5	-.204548	-.038317	-.020637	-.198971	-.037327	-.019227
6	-.227652	-.070841	-.049063	-.242448	-.069749	-.046373

We performed EFA for the six GM and six corresponding EM survey items, respectively, using the half of the dataset we designated for EFA. Models were tested for one and two factors for each survey, each with four patterns of correlated uniqueness: zero CUs specified; one CU specified, Item 5 with 6; two CUs specified, Item 1 with 4 and Item 2 with 3; and all three CUs specified. We decided to include both Item 1 with 4 and Item 2 with 3 in any model that contained one of the CUs because the logic identifying them as similar was the same for these item pairs—both had the same phrasing at the end of the item. Fit indices for each model are given in Table 6; those below rule-of-thumb thresholds are bolded.

Table 6. Fit indices for EFA models, GM and EM

Survey	# Factors	CU	# Parameters	χ^2	df	CFI	RMSEA	SRMR
GM	1	0	24	**127.983**[d]	9	.900	.114	**.050**
EM	1	0	24	**156.890**[d]	9	.896	.129	**.057**
GM	1	1	25	**105.670**[d]	8	.918	.109	**.044**
EM	1	1	25	**90.082**[d]	8	.942	.102	**.039**
GM	1	2	26	**41.912**[d]	7	**.971**	.070	**.031**
EM	1	2	26	63.263	7	**.960**	.090	**.038**
GM	1	3	27	**21.519**[b]	6	**.987**	**.050**	**.021**
EM	1	3	27	**13.061**[a]	6	**.995**	**.034**	**.016**
GM*	2	0	29	**30.921**[d]	4	**.977**	.081	**.026**
EM*	2	0	29	**79.973**[d]	4	.947	.138	**.038**
GM	2	1	30	**9.089**[a]	3	**.995**	**.045**	**.013**
EM	2	1	30	2.566	3	**1.000**	**.000**	**.007**
GM*	2	2	31	5.630	2	**.997**	**.042**	**.009**
EM	2	2	31	**6.666**[a]	2	**.997**	**.049**	**.010**
GM*	2	3	32	**5.291**[a]	1	**.996**	.065	**.009**
EM	2	3	32	5.549	1	**.997**	.068	**.010**

Note: χ^2 tests of model fit were significant at [a]$p<.05$; [b]$p<.01$; [c]$p<.001$; [d]$p<.0001$.
*The residual covariance matrix is not positive definite.

We examined all models having at least three fit statistics below the threshold with positive definite covariance matrices for interpretability, including 1-factor models with two and three CUs, and the 2-factor models with one, two, and three CUs. One-factor models were readily interpreted as demonstrating the "Mindset" construct. The two-factor models were more difficult to interpret, with all or all but one items loading onto one or both of the factors; none of the models matched the theorized two-latent variable model of "Smartness" and "Failure." The best model in terms of both interpretability and fit statistics was the one-factor, three-CU model for both the General and Engineering surveys. Parameter estimates for these models can be found in Table 7, below. Note that the parameter estimates for the EM survey are of roughly the same magnitude but opposite sign to those for the GM survey, with the possible exception of Item 5; despite the opposite signs, this indicates that the factor structures are roughly congruent. The single factor is henceforth named "Mindset."

Table 7. EFA Parameter estimates for GM and EM 1-factor models

	GM			EM		
Parameter	Parameter Estimate	Standard Error	P-Value	Parameter Estimate	Standard Error	P-Value
Mindset BY						
Item 1	-0.387	0.045	0.000	0.502	0.044	0.000
Item 2	0.798	0.073	0.000	-0.731	0.047	0.000
Item 3	0.725	0.068	0.000	-0.621	0.048	0.000
Item 4	-0.439	0.047	0.000	0.563	0.044	0.000
Item 5	-0.297	0.045	0.000	0.605	0.046	0.000
Item 6	0.353	0.045	0.000	-0.274	0.046	0.000
Item 1 WITH						
Item 4	0.271	0.040	0.000	0.182	0.040	0.000
Item 2 WITH						
Item 3	0.011	0.096	0.906	0.198	0.055	0.000
Item 5 WITH						
Item 6	-0.155	0.038	0.000	-0.262	0.041	0.000

Confirmatory Factor Analysis

Our next step was to use CFA with the non-EFA dataset to confirm the suitability of the chosen EFA model, with six items loading onto one factor, and three CUs specified. To avoid having an under identified model, one parameter for each model needed to be fixed. For each of the GM and EM surveys, therefore, we fitted two factor models: 1) in Model A, the factor variance was fixed at 1 and all other parameters were freed; and 2) in Model B, the item with the highest parameter estimate, Item G2 (on the General Mindset survey), was fixed at .798, and the factor variance was freed. We chose a single, positive value for Item 2 for both surveys, forcing the EM survey factor scores to be on a positive scale (larger numbers indicating a growth mindset). We have included parameter estimates for these models in Tables 9 and 10, and fit indices (identical for Models A and B) in Table 8.

Table 8. Fit indices for CFA models, GM and EM

Survey	# Parameters	χ^2	df	CFI	RMSEA	RMSEA 95% CI	SRMR
GM	27	7.875	6	**.998**	**.017**	**.000-.046**	**.014**
EM	27	**13.457**	6	**.992**	**.036**	**.008-**.061	**.018**

Table 9. CFA Parameter estimates for GM models A and B

Parameter	Model A			Model B			R-Square	Residual Variance
	Estimate	S.E	P-Value	Estimate	S.E.	P-Value		
Mindset BY								
Item G1	-0.387	0.054	0.000	-0.461	0.099	0.000	0.150	0.850
Item G2	0.670	0.076	0.000	0.798	N/A	N/A	0.449	0.551
Item G3	0.547	0.071	0.000	0.651	0.068	0.000	0.299	0.701
Item G4	-0.421	0.056	0.000	-0.501	0.105	0.000	0.177	0.823
Item G1 WITH								
Item G4	0.284	0.043	0.000	0.284	0.043	0.000		
Item G2 WITH								
Item G3	0.239	0.080	0.003	0.239	0.080	0.003		
Item G5 WITH								
Item G6	-0.120	0.040	0.003	-0.120	0.040	0.003		
Variance of								
Mindset	1.000	N/A	N/A	0.706	0.161	0.000		

Table 10. CFA Parameter estimates for EM models A and B

Parameter	Model A			Model B			R-Square	Residual Variance
	Estimate	S.E	P-Value	Estimate	S.E.	P-Value		
Mindset BY								
Item E1	0.371	0.047	0.000	-0.388	0.069	0.000	0.137	0.863
Item E2	-0.763	0.065	0.000	0.798	N/A	N/A	0.583	0.417
Item E3	-0.678	0.062	0.000	0.709	0.060	0.000	0.460	0.540
Item E4	0.494	0.046	0.000	-0.516	0.077	0.000	0.244	0.756
Item E5	0.468	0.051	0.000	-0.489	0.078	0.000	0.219	0.781
Item E6	-0.338	0.047	0.000	0.354	0.063	0.000	0.114	0.886
Item E1 WITH								
Item E4	0.324	0.039	0.000	0.324	0.039	0.000		
Item E2 WITH								
Item E3	0.018	0.076	0.810	0.018	0.076	0.810		
Item E5 WITH								
Item E6	-0.100	0.049	0.041	-0.100	0.049	0.041		
Variance of								
Mindset	1.000	N/A	N/A	.915	.155	0.000		

Both the General and Engineering models show good fit. The variation between the corresponding parameter estimates for Models A and B, however, is enough to warrant further investigation. Using the full samples, we fit both models again for each survey to examine the parameter estimates for Item 2 and for the variance. For Model A, the General Mindset estimate for Item 2 was 0.736; the Engineering Mindset estimate for Item 2 was -0.735. Variance was slightly different between the survey models (General estimate = 0.995; Engineering estimate = 1.002). Given these findings and the preference for both Mindset factor scores to have the same direction of factor loadings for similar interpretation, we decided that the factor scores would be built from

Model B, with the parameter for Item 2 fixed at 0.736. We then output factor scores for each student in the full datasets for further analysis using Mplus. Final parameters for the output factors are given in Table 11.

Table 11. Parameters for GM and EM output factor scores

	GM			EM		
Observed item	Parameter	S.E.	P-Value	Parameter	S.E.	P-Value
Mindset BY						
Item 1	-0.388	0.055	0.000	-0.443	0.046	0.000
Item 2	0.736	N/A	N/A	0.736	N/A	N/A
Item 3	0.635	0.040	0.000	0.641	0.035	0.000
Item 4	-0.437	0.060	0.000	-0.534	0.050	0.000
Item 5	-0.320	0.046	0.000	-0.543	0.050	0.000
Item 6	0.326	0.046	0.000	0.304	0.038	0.000
Item 1 WITH						
Item 4	0.275	0.029	0.000	0.249	0.028	0.000
Item 2 WITH						
Item 3	0.133	0.063	0.035	0.122	0.044	0.005
Item 5 WITH						
Item 6	-0.137	0.028	0.000	-0.180	0.032	0.000
Variance of						
Mindset	0.995	0.144	0.000	0.999	0.105	0.000

Test Information Functions

Test information functions and the standard error of measurement (SEM) are given in Figure 2. These curves show that the amount of information available is reasonable (SEM approximately 0.5) for scores at or below the mean, but drops off for mindset scores above the mean, with scores that are a standard deviation above the mean determined with relatively low information and high SEM (Samajima, 1994).

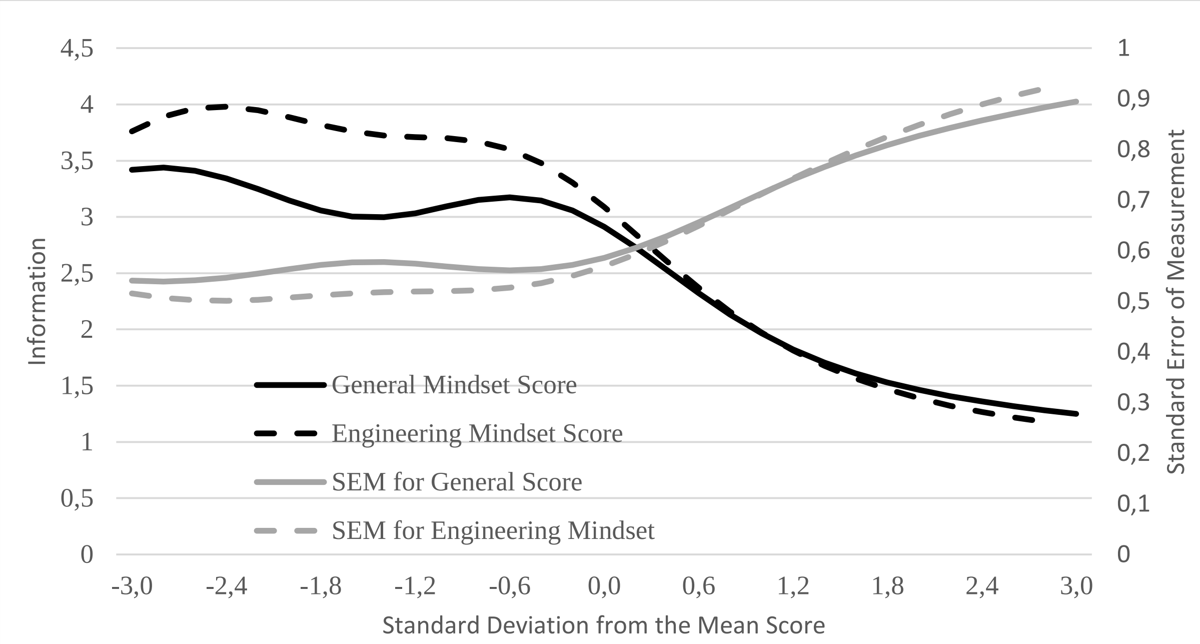

Figure 2. Test information functions showing information by standard deviations from the mean score

Examination of the item information curves for each survey shows that Item 6 contributes the least information overall to either score (see Figures 3 & 4). Item 5 is a good contributor to the EM but not the GM score. Items 2 and 3, meanwhile, are the best contributors to both scores.

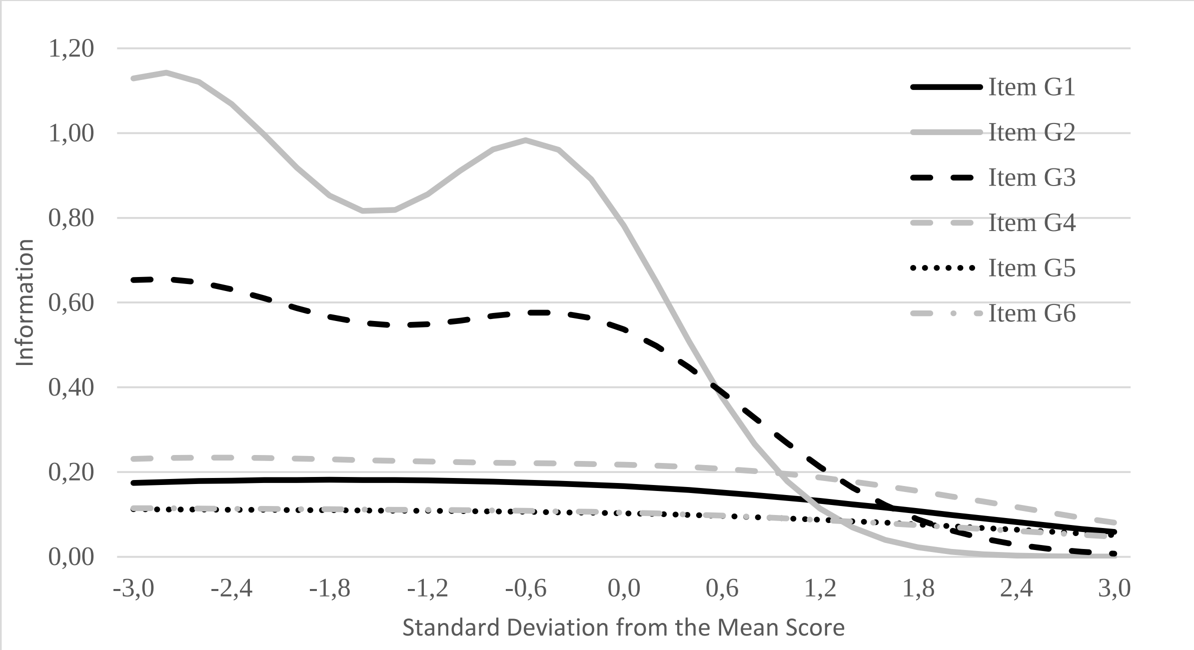

Figure 3. Item information curves showing information by standard deviations from the mean GM score

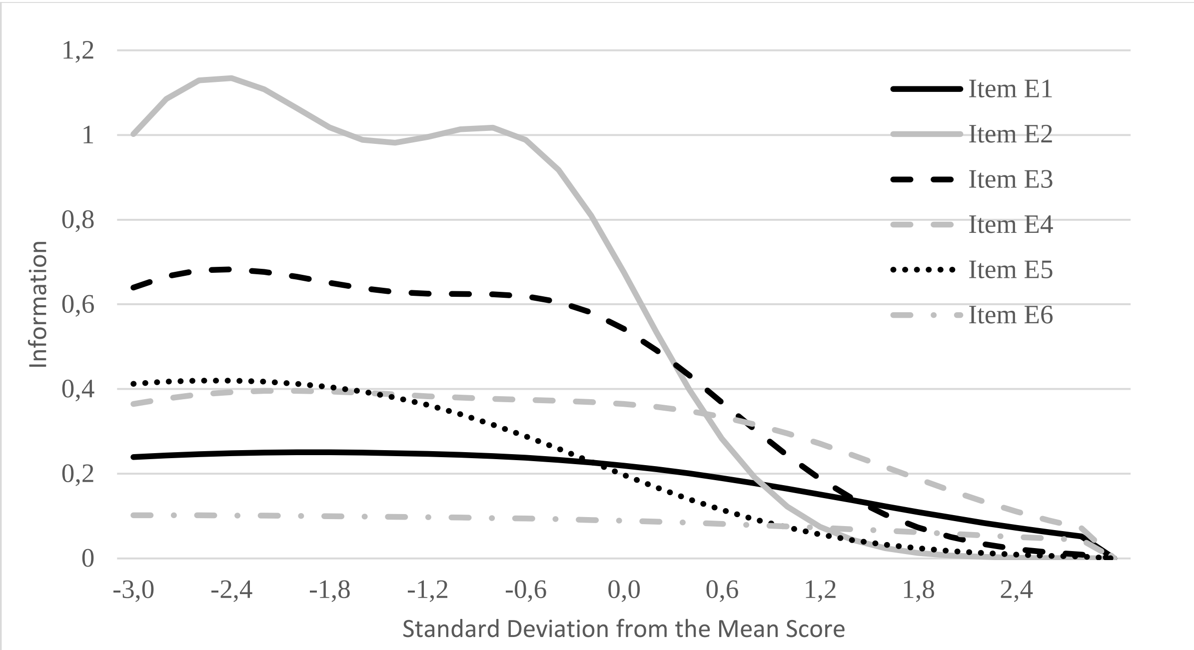

Figure 4. Item information curves showing information by standard deviations from the mean EM score

Conclusions

As discussed previously, various instruments have been used to measure mindset of K12 students (Dweck, 1999; Blackwell et al., 2007; Park et al., 2016; PERTS, 2015; Petscher et al., 2017; Stipek & Gralinski, 1996). Our task was to develop GM and EM surveys that built upon both Dweck's (1999) survey and our own expertise with respect to failure in the context of engineering (Lottero-Perdue 2015; Lottero-Perdue & Parry, 2017a, 2017b).

The surveys we describe here each show evidence that the items describe one latent factor, mindset. All six items load significantly on their respective factors. In addition to evidence of structural validity, we demonstrate significant evidence of content validity through the strong foundation of item content in prior research, and in qualitative interviews and testing with upper elementary students. Based on this evidence we have collected, we assert that there is strong evidence for the validity of using either the GM or EM survey with students ages 10-11 who have engaged in engineering classwork in school. We expect that other researchers will use these surveys for further research with students of similar ages.

This research contributes a new, evidence-based pair of instruments for use in examining the mindset of students aged 10 and 11 years. The EM survey in particular is intended for use in examining the mindset of young students engaged in engineering learning. There is a need for domain-specific survey instruments to compare general mindset to domain-specific mindset in the currently active research program around students' growth versus fixed mindset and the relationship between mindset and student learning outcomes and achievement. This paper contributes a domain-specific survey instrument for engineering. It also sheds further light on the concept of failure, demonstrating that failure has a strong relationship to the mindset construct.

Limitations and Future Research

The GM and EM surveys were designed for fifth-grade students aged 10 and 11 years. The sample of students we used to examine reliability and provide evidence for validity was a relatively diverse sample of students with respect to gender, race/ethnicity, and socioeconomic status, and includes those who receive special education services. It also includes English Language Learners, albeit we have chosen not to present these data given the high percentage (94%) of missing data regarding English proficiency.

Although we did not translate the GM and EM surveys into other languages such as Spanish, we suggest that future administration of the surveys include this as an option for English Language Learners (Lachapelle & Brennan, 2018). We would also suggest including specific instructions regarding possible accommodations for students with special needs, including reading items aloud and providing a more visual way to respond to the Likert questions as done by Gunderson and colleagues (2013) and others for younger students.

We suspect that the surveys could be used with success at the middle school level and with slightly younger – e.g., fourth-grade – students. Language about "smartness" may seem juvenile to those at the high school level and the phrasing of the items may be abstract for those in the younger grades. Researchers and educators desiring to use these surveys with students younger or older than our sample population should collect further evidence of validity and reliability.

As explained previously, we administered the GM survey prior to engineering instruction and the EM survey after engineering instruction. Pilot testing with a combined set of both GM and EM items had yielded feedback from students that items on the survey seemed redundant. That said, surveys that include both general and domain-specific items have been used in the mindset research (e.g., Petscher et al., 2017). Regardless of whether the surveys are used together at one point in time or separated, our major recommendation is to use the EM survey only after students have learned about engineering. Providing the EM survey before such instruction would yield results that would be difficult to interpret given that many elementary students are unfamiliar with engineering. Likewise, we do not recommend that the EM survey be used as a pre-post survey, with EM before and then EM after instruction in an attempt to ascertain growth in mindset, without further investigation as well as consideration of the prior engineering knowledge of students.

We intend to use the results from GM and EM surveys in our future research. Specifically, we will investigate how students' GM impacts their interests in and attitudes about engineering and their performance on engineering assessments. A strong relationship between GM and these outcomes may suggest the use of growth mindset interventions within engineering instruction. As discussed earlier, such interventions have produced gains in other subject areas for older students (e.g., Blackwell, Trzesniewski, and Dweck, 2007). Further, we are interested in the relationship between demographic characteristics and mindset, which may inform what particular groups may benefit the most from growth mindset interventions.

Others may wish to use the GM and EM surveys to explore students' mindsets and to develop and implement ways to grow growth mindsets. One question researchers may aim to explore through the use of these surveys is: How does EM change with time? For example, do students' mindsets about engineering become more fixed as they move from elementary to middle school and then from middle school to high school given engineering education experience at each of these levels? A related question: Is this pattern the same for those who are represented in engineering (e.g., those who are white, those who are male, etc.) and those who are not represented in engineering (e.g., those who are black, those who are Hispanic, those who are female). Another way to use the GM and EM surveys is to measure mindset to determine the extent to which mindset-related interventions could be employed to support the development of growth mindsets in general or with respect to engineering among students.

While we have found evidence to support the validity and reliability of the GM and EM surveys, there is room for improvement. As mentioned previously, reliability of the factor scores for each of the surveys drops substantially for scores above the mean and is unreliable for scores at a standard deviation or more above the mean. This suggests that it would be wise to develop new survey questions that can more reliably and precisely identify students with the most flexible mindsets for both surveys. Beyond this, Items 3 and 4 on each survey seem important to continue to include on both GM and EM surveys. However, while it may be useful to include failure items (Items 5 and 6) on EM surveys – perhaps given the unique role of failure within engineering – it may not be necessary to include them on GM surveys.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1220305. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Adelson, J.L. & McCoach, D.B. (2010). Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point Likert-type scale. *Educational and Psychological Measurement, 70*(5) 796-807. Doi: 10.1177/00131644103666694

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences*. New York, NY: Routledge.

Blackwell, L.S., Trzesniewski, K.H., & Dweck, C. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. http://www.jstor.org/stable/4139223

Brown, T. A. (2006). *Confirmatory factor analysis for applied research* (1st ed.). New York, NY: The Guilford Press.

Crismond, D.P. & Adams, R.S. (2012). The Informed Design Teaching and Learning Matrix. *Journal of Engineering Education, 101*(4), 738-797. doi: 10.1002/j.2168-9830.2012.tb01127.x

Cunningham, C. (2018). *Engineering in elementary STEM education: Curriculum design, instruction, learning and assessment.* New York: Teachers College Press.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, *14*(20), 1–11.

Douglas, K. A., & Purzer, Ş. (2015). Validity: Meaning and relevancy in assessment for engineering education research. *Journal of Engineering Education*, *104*(2), 108–118. https://doi.org/10.1002/jee.20070

Duncan, O. D., Featherman, D. L., & Duncan, B. (1972). *Socioeconomic background and achievement.* Oxford, England: Seminar Press.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 41*(10), 1040-1048. doi:10.1037/0003-066X.41.10.1040

Dweck, C.S. (1999). *Self-theories: their role in motivation, personality and development.* Philadelphia, PA: Psychology Press.

Dweck, C.S. (2006). *Mindset: The New Psychology of Success*. New York, NY: Random House.

Dweck, C.S. (2015, September 22). Carol Dweck revisits the 'Growth Mindset.' *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2015/09/23carol-dweck-revisits-the-growth-mindset.html

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, *16*(1), 5.

Engineering is Elementary. (2019). The Engineering Design Process. Retrieved March 21, 2019 from http://www.eie.org/overview/engineering-design-process

Good, C., Rattan, A., & Dweck, C.S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology, 102*(4), 700-717. doi:10.1037/a0026659

Good, C., Aronson, J. & Inzlicht, M. (2003). Improving adolescent's standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*(6), 645-663. doi:10.1016/j.appdev.2003.09.002

Gunderson, E.A., Sorhagen, N.S., Gripshover, S.J., Dweck, C.S., Goldin-Meadow, S., & Levine, S.C. (2018). Parent praise to toddlers predicts fourth-grade academic achievement via children's incremental mindsets. *Developmental Psychology, 54*(3), 397-409.

Hanson, J. (2017). Determination and validation of the Project for Educational Research that Scales (PERTS) survey factor structure. *Journal of Educational Issues, 3*(1), 64-82. Doi: 10.5296/jei.v3i1.10646

Heyman, G.D.. & Dweck, C.S. (1998). Children's thinking about traits: Implications for judgemtns of the self and others. *Child Development, 69*, 391-403.

IBM Corporation. (2016). IBM SPSS Statistics for Windows (Version 24). Armonk, NY: IBM.

Kelloway, E. K. (2015). *Using Mplus for structural equation modeling: A researcher's guide* (2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Lachapelle, C. P. & Brennan, R. T. (2018). An instrument for examining elementary engineering student interests and attitudes. *International Journal of Education in Mathematics, Science and Technology, 6*(3), 221-240. doi:10.18404/ijemst.428171

Lottero-Perdue, P.S. (2015, April). The Engineering design process as a safe place to try again: Responses to failure by elementary teachers and students. Presented at the NARST Annual Conference, April 11, 2015, Chicago, IL.

Lottero-Perdue, P.S. & Parry, E. A. (2017a). Elementary teachers' reflections on design failures and use of fail words after teaching engineering for two years. *Journal of Pre-College Engineering Education, 7*(1), Article 1, 1-24.

Lottero-Perdue, P.S. & Parry, E. A. (2017b). Perspectives on failure in the classroom by elementary teachers new to teaching engineering. *Journal of Pre-College Engineering Education, 7*(1), Article 4, 1-21.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*(4), 810.

Matheson, I. (2015). Self-regulatory efficacy and mindset of at-risk students: An exploratory study. *Exceptionality Education International, 25*(1), 67-90. Retrieved from http://ir.lib.uwo.ca/eei/vol25/iss1/4

McCutchen, K.L., Jones, M.H., Carbonneau, K.J., & Mueller, C.E. (2016). Mindset and standardized testing over time. *Learning and Individual Differences, 45*. 208-213. Doi:10/1016/j.lindif.2015.11.027

Moore, T.J., Glancy, A.W., Tank, K.M., Kersten, J.A., Smith, K.A., Stohlmann, M.S. (2014). A framework for quality K-12 engineering education: Research and development. *Journal of Pre-College Engineering Education Research, 4*, Article 1.

Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus user's guide*. (8th Ed.). Los Angeles, CA: Muthén & Muthén

National Center for Education Statistics. (2012). Improving the measurement of socioeconomic status for the National Assessment of Educational Progress: A theoretical foundation (p. 37). Washington, D. C.: National Center for Education Statistics. Retrieved from http://nces.ed.gov/nationsreportcard/researchcenter/socioeconomic_factors.aspx

Next Generation Science Standards (NGSS) Lead States (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press. doi:10.17226/18290

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, *32*(3), 396–402. https://doi.org/10.3758/BF03200807

Paunesku, D., Walton, G.M., Romero, C., Smith, E.N., Yeager, D.S. & Dweck, C.S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793.

Park, D., Gunderson, E.A., Tsukayama, E., Levine, S.C., & Beilock, S.L. (2016). *Journal of Educational Psychology, 108*(3), 300-313. Doi: 10.1037/edu0000064

PERTS (Project for Education Research that Scales) (2015). Growth mindset assessment. Retrieved from https://survey.perts.net/share/toi

Petscher, Y., Al Otiaba, S., Wanzek, J., Rivas, B., & Jones, F. (2017). The relation between global and specific mindset with reading outcomes for elementary school students. *Scientific Studies of Reading, 21*(5), 1-16. doi: 10.1080/10888438.2017.1313846

Ricci, M.C. (2013). *Mindsets in the Classroom: Building a Culture of Success and Student Achievement in Schools*. Prufrock Press, Inc.: Waco, TX.

Ruel, E., Wagner, W.E. III, & Gillespie, B.J. (2016). *The practice of survey research: Theory and applications*. Sage Publications, Inc.: Thousand Oaks, CA.

Samajima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement*, *18*(3), 229–244.

Schroder, H.S., Fisher, M.E., Lin, Y., Lo, S.L., Danovitch, J.H., & Moser, J.S. (2017). Neural evidence for enhanced attention to mistakes among school-aged children with a growth mindset. *Developmental Cognitive Neuroscience, 24*, 42-50. doi: 10.1016/j.dcn.2017.01.004

Sijtsma, K. (2009). On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453.

Stipek, D. & Gralinski, J.H. (1996). Children's beliefs about intelligence and school performance. *Journal of Educational Psychology, 88*(3), 397-407. doi: 10.1037/0022-0663.88.3.397

Tomasetto, C., Alparone, F.R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology, 47*(4), 943-949. Doi: 10.1037/a0024047

Wonch Hill, P., McQuillan, J., Talbert, E., Spiegel, A., Gauthier, G., Diamond, J. (2017). Science possible selves and the desire to be a scientist: Mindsets, gender bias, and confidence during early adolescence. *Social Sciences, 6*(2), 55. doi:10.3390/socsci6020055

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., & ... Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*(3), 374-391. doi:10.1037/edu0000098

Author Information

Pamela S. Lottero-Perdue
Department of Physics, Astronomy & Geosciences
Towson University
8000 York Road, Towson, MD 21252
USA
Contact e-mail: *plottero@towson.edu*

Cathy P. Lachapelle
Museum of Science, Boston
Engineering is Elementary
1 Science Park, Boston, MA 02114
USA

Appendix: Surveys

General Mindset Survey: We are interested in how you think about being intelligent (or smart). We also want to know what you think it means if you try something and fail (or do not succeed) at it.

Please check how much you agree or disagree with each statement:

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. You can't really change how smart you are.	☐	☐	☐	☐
2. No matter how smart you are now, you can always become a lot smarter.	☐	☐	☐	☐
3. No matter who you are, you can become a lot smarter.	☐	☐	☐	☐
4. You can learn new things, but you can't really change how smart you are.	☐	☐	☐	☐
5. If you try and fail at something, that means you are not smart at that kind of thing.	☐	☐	☐	☐
6. If you try and fail at something, that means that you would want to try to do that thing again.	☐	☐	☐	☐

Engineering Mindset Survey: We are interested in how you think about being a smart (or a good) engineer. We also want to know what you think it means if you create a design that fails.

Please check how much you agree or disagree with each statement:

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. You can't really change how smart you are at engineering.	☐	☐	☐	☐
2. No matter how smart you are at engineering now, you can always become a lot smarter.	☐	☐	☐	☐
3. No matter who you are, you can become a lot smarter at engineering.	☐	☐	☐	☐
4. You can learn new things about engineering, but you can't really change how smart you are at engineering.	☐	☐	☐	☐
5. If your design fails, that means you are not smart at engineering.	☐	☐	☐	☐
6. If your design fails, that means that you would want to engineer a new design.	☐	☐	☐	☐

Questions 1-4 were on each survey adapted from Carol Dweck's mindset surveys (1999, 2006).