122nd ASEE Annual Conference & Exposition
June 14 - 17, 2015
Seattle, WA
Paper ID #11925

The Four-Phase Interest Development in Engineering Survey

Joseph E Michaelis, University of Wisconsin - Madison

Joseph E Michaelis is a Ph.D. student in Educational Psychology in the Learning Sciences area at the University of Wisconsin - Madison. His research involves studying interest in STEM education, focusing on the impact of learning environments, feedback, and influence of social constructs and identities. This research includes developing inclusive learning environments that promote interest in pursuing STEM fields as a career to a broad range of students.

Prof. Mitchell Nathan, University of Wisconsin, Madison

Mitchell J. Nathan is a professor of learning sciences in the Department of Educational Psychology at the University of Wisconsin-Madison. Prof. Nathan received his Ph.D. in experimental (cognitive) psychology. He also holds a B.S. in electrical and computer engineering, mathematics and history. He has worked in research and development in artificial intelligence, computer vision and robotic mobility, including: design and development of autonomous robotic arms and vehicles; sensor fusion; the development of expert systems and knowledge engineering interview techniques; and the representation of perceptual and real-world knowledge to support inference-making in dynamic environments. Nathan also has worked on computer-based tutoring environments for mathematics education that rely heavily on students' own comprehension processes for self-evaluation and self-directed learning (so-called unintelligent tutoring systems). Prof. Nathan has authored over 100 peer-reviewed publications, given more than 120 presentations at professional meetings, and has secured over $25M in research funds to investigate and improve STEM learning, reasoning and instruction. Among his projects, Dr. Nathan directed the IERI-funded STAAR Project, which studied the transition from arithmetic to algebraic reasoning, served as Co-PI for the NSF-funded AWAKEN Project, which documented how people learn and use engineering, and currently co-directs the National Center for Cognition and Mathematics Instruction. He is a faculty member for the Latin American School for Education, Cognitive and Neural Sciences. As part of his service to the nation, Dr. Nathan served on the National Academy of Engineering/National Research Council Committee on Integrated STEM Education, and is currently a planning committee member for the Space Studies Board of the National Academy of Sciences/National Research Council workshop Sharing the Adventure with the Student: Exploring the Intersections of NASA Space Science and Education. At the University of Wisconsin, Dr. Nathan holds affiliate appointments in the Department of Curriculum & Instruction, the Department of Psychology, and the Wisconsin Center for Education Research. He is a member of the steering committee for the Delta Program (part of the national CIRTL Network), which promotes the development of a future national STEM faculty committed to implementing and advancing effective teaching practices for diverse student audiences. Prof. Nathan currently is Director of the Center on Education and Work and Director of the Postdoctoral Training Program in Mathematical Thinking, Learning, and Instruction. He is an inductee and executive board member of the University of Wisconsin's Teaching Academy, which promotes excellence in teaching in higher education.

The Four-Phase Interest Development in Engineering Survey

Abstract

The Four-Phase Interest in Engineering Survey (FIDES) was developed to address a need for a survey instrument derived from current interest development theory to measure the psychological construct of student interest in engineering. The FIDES survey was developed using the Four-Phase Model of Interest Development as a theoretical framework for identifying indicators of student interest. In this paper we discuss the design and development of the FIDES instrument using an iterative method. After two rounds of modification and fine tuning, we present the FIDES 2.0 instrument as a valid and reliable tool for measuring student interest in engineering. We explore the merits and limitations for the FIDES 2.0 to inform our understanding of interest as a developing psychological construct, and suggest appropriate applications that would benefit from using the survey.

Introduction

Increasing motivation for students to enter and persist in engineering pathways will require developing and maintaining interest in engineering[1]. In order to better estimate the effectiveness of programs aimed at increasing this interest, a measurement instrument based in interest development theory is needed[2]. Use of a survey instrument will provide a standardized and efficient tool for this purpose. Currently, most available survey instruments that measure student interest in engineering are based on affective measures of interest[3], or attitudes towards STEM[4], without accounting for indicators of deeper and more meaningful interest. It is important to include indicators that go beyond the surface level of positive feelings in order to better account for a more nuanced understanding of interest development. One currently used survey, which includes measures beyond affect and attitude, was implemented by Maltese, Melki, and Wiebke[5]. However the instrument utilizes a retrospective approach, that is not suitable for measuring current interest in engineering, and has the drawback that retrospective recall of mental states are notoriously inaccurate[6]. The goal of this study is to develop and refine a psychometrically sound survey instrument based on Hidi and Renninger's[7] Four-Phase Model of Interest Development (FPMID) that will reliably ($\alpha \geq .70$) measure students' early and developing levels of interest in engineering.

Background Literature and Theoretical Framework

Education theorists distinguish between state-based and trait-based personality characteristics of students. The *traits* of a student "are relatively constant over time" regardless of variations between situations and task demands (though they will vary between students), while the *states* that a student goes through can frequently "vary during individual learning experiences" [8]. An example of a state would be the changing nature of a student's content knowledge over the course of a lesson. The student's mental state can change as she learns a new concept or procedure, which can lead her to think and act differently in the future. An example trait is a student's level of extraversion, and will likely remain unchanged as she moves from class to class, or even as she continues to mature.

Interest as a developing psychological construct may appear to the observer initially as a *state* for a student whose interest in computer programming is only triggered by engaging situations, such as an engaging Lego/Logo activity. Over time, however, it may become more evident that programming interest is a *trait* for this student, who personalizes and internalizes a drive to explore and pursue a degree in electrical and computer engineering. Previous work has identified an important distinction between *situational interest*, which is state-based, context-specific, impulsive, and transitory; and *individual interest*, or personal interest, which is trait-based, context independent and enduring[9, 10, 11]. Hidi and Renninger's[7] FPMID further refines the situational-individual interest distinction into four progressive phases of interest: Triggered Situational, Maintained Situational, Emerging Individual, and Well-Developed Individual. Each phase is considered to be sequential and distinct, with each subsequent phase building off the previous. Triggered and Maintained Situational phases of interest are hypothesized to be primarily state-based, while Emerging and Well-Developed Individual phases are considered to be trait based. Over time, and through repeated activation, states can develop into traits, through neural reorganization during brain development[12]. This is one reason why early experiences that first "catch" and then "hold" one's interest are thought to have such a sustained effect on later interest development[13, 14].

Hidi and Renninger's model provides empirically driven descriptive characteristics of students in each phase of interest (see Table 1). These descriptive characteristics allow insight into measurable *indicators* of interest that go beyond surface level descriptors like enjoyment and positive affect, and are thus well suited for measuring interest in a psychometrically reliable manner. Use of these indicators will enable the design of an instrument to be sensitive to students at all phases of interest in engineering, and can be used to assess changes in interest development that is crucial to understanding program impact on student interest in engineering.

Table 1:
Learner characteristics in each of the four phases of interest development (adapted from Renninger, 2009).

Phase	Triggered Situational	Maintained Situational	Emerging Individual	Well-Developed Individual
Content	Attend to content, if only fleetingly	Re-engage content that previously triggered attention	Are likely to independently re-engage content	Independently re-engage content
Support	Need support to engage with others (e.g., group work) and with instructional design (e.g., software)	Supported by others to find connections between their knowledge, skill, and prior experience	Pose curiosity questions that lead them to seek answers Focused on their own questions	Pose curiosity questions Self-regulates easily to reframe questions and seek answers
Feelings	Experience either positive or negative feelings	Positive feelings	Positive feelings	Positive feelings
Knowledge		Developing knowledge of the content	Have stored knowledge	Have stored knowledge
Personal Value		Developing a sense of the content's value	Have stored value	Have stored value
Feedback	Want to simply be told how to complete assigned tasks in as few steps as possible	Want to be told what to do	May have little value for the canon of the discipline and most feedback	Actively seeks feedback Recognize others' contributions to the discipline
Perseverance				Can persevere through frustration and challenge in order to meet goals

In an effort to address the goal of a reliable means for measuring student interest based on the indicators of the FPMID, we report here on the methodological process for developing and refining an instrument we term the Four-Phase Interest Development in Engineering Survey (FIDES). As we show, FIDES has the potential to enable future studies to measure pre- and post-participation levels of interest in engineering for high school students. Analyses of findings from

this instrument suggest that the FIDES can be an effective means of measuring program impact on student interest development over time, as well as to document individual differences.

Method

The method of instrument development we used relies on an iterative process. Initial instrument construction starts with theoretically motivated items that capture the hypothesized dimensions of engineering interest. The first version of the survey (FIDES 1.0) is then administered to the target population (n = 197), and that initial data provides us with an empirical basis to assess and improve instrument reliability and validity. Several insights from that analysis led us to develop a revised version, FIDES 2.0, which was then administered to a second sample in the target population (n = 145).

FIDES 1.0: Initial Instrument Development

Item Construction

An initial list of six indicators of interest from Hidi and Renninger's FPMID were identified and used to create instrument items for the FIDES 1.0 survey instrument. The indicators include: independent content engagement (CONTENT), independent questioning (QUES), positive feelings (FEEL), use of feedback (FEED), perseverance (PERSEVERE), and content knowledge (KNOW). Four of the indicators (QUES, FEEL, FEED, PERSEVERE) were assessed using two positively phrased Likert scale items and one negatively phrased Likert scale item, which asked students to rate their level of agreement with statements using a scale from 1 (strongly disagree) to 7 (strongly agree). The fifth indicator, CONTENT, was assessed using the same Likert scale, but included four positively phrased and two negatively phrased items. This indicator had additional questions in order to include items that expressed multiple ways of engaging with content. A potential sixth indicator (KNOW) was not included on the initial survey due to concerns over the reliability of measuring content knowledge levels using the self-reporting method of a survey. A list of the initial 18 items developed for the study is available in Table 2. To guide participants' while responding to statements about their interest in engineering, we included the following statement in the survey directions to clarify what we intended:

> In the next section you will be asked a series of questions about how you feel about ENGINEERING. This includes any ideas, problems or projects that have to do with engineering. If it helps to make a decision about how to answer, think of an engineering subject you are familiar with or the ideas and topics in engineering you know about when answering."

Table 2: FIDES 1.0 Instrument Items

Item Label	Statement
CONTENT 1	I find things about engineering interesting when I'm outside of school.
CONTENT 2	When I find something to be interesting in engineering, I pay greater attention overall in science class.
CONTENT 3	When I find something to be interesting in engineering, I love to

	learn more about it when I'm outside of class.
CONTENT 4	When I find something to be interesting in engineering, I continue thinking about that topic for more than two weeks.
CONTENT 5 – NEG*	When I find something to be interesting in engineering, I NEVER think about it outside of class.
CONTENT 6 – NEG*	When I find something to be interesting in engineering, I NEVER learn more about it outside of class.
QUES 1	When I find something to be interesting in engineering, I think of my own questions about that topic.
QUES 2	I like to think of my own questions in engineering.
QUES 3 – NEG*	When I find something to be interesting in engineering I DO NOT like to find my own answers to questions about that topic.
FEEL 1	Knowing about engineering is very valuable.
FEEL 2	Knowing about engineering is very useful.
FEEL 3 – NEG*	Knowing about engineering is useless.
FEED 1	I seek out information from teachers on how to answer questions in engineering.
FEED 2	When I'm working on something I find interesting in engineering, I like to get constructive criticism about how to do better.
FEED 3 – NEG*	I find it FRUSTRATING when I receive criticism about how to improve when I'm working on engineering ideas and concepts.
PERSEVERE 1	I like to learn about ENGINEERING TOPICS even when they are very difficult.
PERSEVERE 2	When I find something to be interesting in engineering, I continue working on that topic even when it becomes very difficult.
PERSEVERE 3 – NEG*	I STOP working on topics in engineering when it becomes difficult.

* Items that were subsequently removed from FIDES 1.0 instrument.

Development of a trustworthy survey to identify the underlying psychological constructs hypothesized by the FPMID depends on two psychometric parameters: reliability and validity. *Reliability* metrics specify the degree to which the individual items of the survey consistently

measure the intended construct (internal consistency). Reliability is typically measured using Cronbach's alpha, with an alpha value of 0.7 or greater indicating, by convention, a reasonably reliable construct.[15] As a further check of the trustworthiness of the instrument at the construct level, we check whether the mean scores of each of the five constructs (e.g., CONTENT, QUES, etc.) each reside near the midpoint of the scale, to avoid construct measures that are not overly skewed.

The second parameter, validity, takes many forms, but they each indicate, in different ways, the degree to which the instrument accurately measured the intended underlying construct. *Content validity* for this survey is exhibited by showing that this instrument reflects all of the dimensions of interest described by the FPMID, including: independent content engagement (CONTENT), independent questioning (QUES), positive feelings (FEEL), use of feedback (FEED), and perseverance (PERSEVERE). As a way to establish *construct validity*, we examine the degree to which all items on the FIDES survey instrument load onto one factor in a confirmatory factor analysis[16]. Another way to establish construct validity is to show convergence between FIDES and an independent instrument for measuring interest. Toward this end, we included an additional seven-item adjective-pair survey, adapted from the Engineering section of the STEM Semantics Survey (E-SSS)[3]. E-SSS is a previously published, valid and reliable survey that is based on an affective interpretation of interest[17, 18]. E-SSS scores are calculated as an average score for the seven items. A high correlation between the E-SSS and FIDES survey would support construct validity for the FIDES instrument. Two modifications were made to the original E-SSS survey items. The first modification moved all positive adjectives in each pair to the right side of the scale, to avoid participant misinterpretation. The second modification was to change the adjective used to anchor the scale from "mundane" to "uninteresting" in order to eliminate any reading comprehension issues for our participants.

As a technical point, FIDES scores are calculated by: 1) first calculating the mean scale score for each indicator, and then 2) averaging a sum of scores from all indicators on the survey to produce scores on the 1 to 7 Likert scale equivalent. FIDES scores here do not represent ordinal phases in the FPMID, as is often done in research using this model. We are presenting a scale score for the interest as a measured construct. To avoid confusion, throughout the paper, we will refer to phases by name rather than phase number. The FIDES scale score represents the equivalent numerical value along the Likert-scale, (from strongly disagree to strongly agree). This value is a measure of the level of agreement that participants report with regard to their interest in engineering.

Additional items were included in the survey administration to collect demographic information about participants, including: gender, grade level, current science course enrollment, current science teacher, and intended college major. Demographic information was collected in order to assess any group differences or confounding variables that may influence instrument results. We expect mean scale scores to show little differences between grade level and current teacher. However, based on prior research that has established a difference of interest in engineering for high school aged males and females[19] we expect the scale to show significant differences between gender. We also anticipate that students indicating an intended college major in a STEM field should also have a significantly higher interest in engineering, as previous work has identified a link between interest in a domain and further pursuit of that domain[20]. If the survey testing results conform to these expectations, this will also provide further evidence for content validity.

Initial Survey Administration

The initial survey consisting of the 18-item FIDES instrument, seven E-SSS items, and five demographic items was administered to 197 students (92 male, 105 female) across all high school grade levels from a small urban charter high school in a large Midwestern city serving a predominantly African-American population (92%), with a high rate of students receiving free or reduced lunches (93%). The school has 42% of its student body meet or exceed expectations on the state standardized test. At this school, all students in each grade were enrolled in the same science content course: Freshmen took a common biology course, sophomores were enrolled in chemistry, juniors in physics, and seniors were enrolled in an environmental science class. The survey was administered in one online session conducted at the school site. Responses were removed based on pre-established criteria for irregular response patterns in which a participant's response to the positively phrased items was the same as their response for the corresponding negatively phrased item for each indicator. Twenty-three such responses were removed from evaluation. The remaining participants (n = 174; male = 75, female = 99) represented a nearly equal distribution across grade levels (9^{th} = 46, 10^{th} = 41, 11^{th} = 44, 12^{th} = 43) with the same distribution of science courses based on grade levels. Students reported having one of six teachers, with some teachers instructing more than one grade level.

Validity and Reliability Analyses

A principal components factor analysis was performed to assess whether all 18 FIDES items loaded on a single factor. The initial analysis showed that 46.8% percent of the variance was explained by one factor, but a second factor explained an additional 16.8% of variance (see Table 3). A two-factor analysis was conducted using all 18 FIDES items. All positively phrased FIDES items loaded at a practically significant level of 0.5[21] within factor 1, with loadings ranging from 0.731 to 0.868. Negatively phrased items did not highly load significantly with factor 1 (range of -.352 to .0199), and all loaded within factor 2, with loadings ranging from 0.541 to 0.761 (see table 4). Since the factor analysis indicated that positively phrased and negatively phrased items loaded on two separate factors, only the 12 positively phrased items were used in calculating the FIDES 1.0 scale score.

Table 3:
Factor analysis of FIDES 1.0 survey, variance explained.

Component	Variance Explained
1	46.8
2	16.8
3	5.1

Table 4:
Factor analysis loadings for FIDES 1.0 items.

Item	Factor	
	1	2
CONTENT 1	.822	.047
CONTENT 2	.816	.090
CONTENT 3	.868	.098
CONTENT 4	.731	.164
CONTENT 5 – NEG	-.352	.719
CONTENT 6 – NEG	-.337	.724
QUES 1	.811	.103
QUES 2	.734	.195
QUES 3 – NEG	-.237	.693
FEEL 1	.818	.032
FEEL 2	.781	.034
FEEL 3 – NEG	-.311	.690
FEED 1	.744	.196
FEED 2	.761	.115
FEED 3 – NEG	.199	.541
PERSEVERE 1	.834	.030
PERSEVERE 2	.821	.082
PERSEVERE 3 – NEG	-.173	.761

An analysis of the 12 remaining positively phrased FIDES items showed high reliability ($\alpha = .94$) for the scale. Scale scores were closely centered on the middle point of the scale ($\bar{X} = 4.2$, $\sigma = 1.4$), and demonstrated a wide range of scores (1.0 to 7.0). The distribution of the scale scores was modeled by a normal distribution ($W = .985$, $p = 0.06$; see Figure 1), had a slight skew to the left (-0.141) and had a severely lower than Gaussian shape (Kurtosis = -0.463). As a way to check construct validity, we performed a correlation analysis of FIDES scores and E-SSS scores. FIDES scores were highly correlated with the E-SSS results ($r = 0.70$).

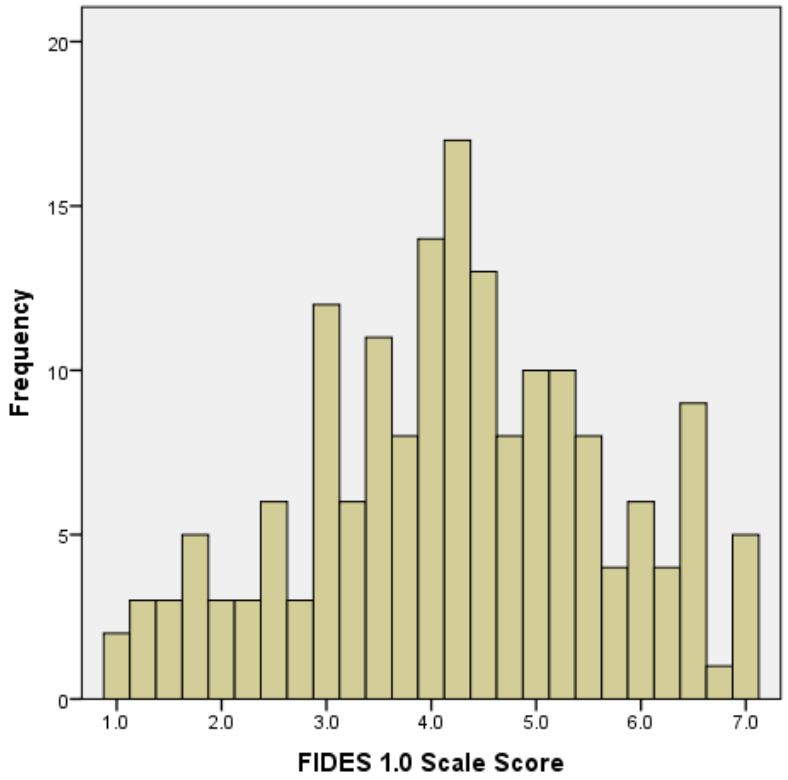

Figure 1: Distribution of FIDES 1.0 scale scores with negatively phrased items removed.

Results

Having established the reliability and the validity of the FIDES 1.0 instrument, we examined how reported interest varied by grade level, teacher, sex, and intended academic major, based on FIDES 1.0 scale scores using the 12 remaining positively phrased items. See Table 5 for scale score means by group.

We found no differences in mean FIDES 1.0 scores based on grade level (F (3,170) = 1.39, p = 0.25), or teacher (F (5,168) = 1.23, p = 0.30). Mean FIDES 1.0 scores for males (4.62) and females (3.94) were found to be significantly different (t (172) = 3.17, $p < 0.05$), showing that male students overall reported higher phases of interest development than female students.

Student reports of intended majors were classified into 5 categories: Undecided (n = 95), STEM fields (n = 41), Humanities (n = 20), Business (n = 9), and Performing Arts (n = 9). An omnibus test for comparison of the means of students based on their intended major revealed a difference in means (F (4,169) = 2.45, p = 0.48). A comparison of the FIDES 1.0 mean for students intending on majoring in a STEM field (4.70) was contrasted with all other intended majors (4.18; not including undecided) and no significant difference was found, (t (77) = 1.51, p = 0.13). Mean scores comparing STEM students with those intending to major in a Humanities field (3.8) indicated a significant difference, (t (169) = 2.31, $p < 0.05$.

Table 5:
Descriptive statistics for mean FIDES 1.0 scores, by demographic grouping

Group	Count	Average	σ
Grade			
9	46	4.17	1.30
10	41	3.82	1.51
11	44	4.44	1.51
12	43	4.37	1.40
Teacher			
A	23	4.01	1.14
B	35	3.79	1.54
C	23	4.33	1.46
D	43	4.43	1.52
E	46	4.42	1.31
F	4	3.45	2.09
Gender			
Male	75	4.58	1.31
Female	99	3.92	1.47
Major			
Undecided	95	4.01	1.32
STEM	41	4.70	1.59
Humanities	20	3.82	1.35
Business	9	4.67	1.87
Performing Arts	9	4.48	1.07

Discussion of Initial Item Testing

Overall, the FIDES 1.0 instrument appears to be a reliable and valid interest measurement instrument. Its high Cronbach's alpha indicates a high internal reliability, while a high correlation with the E-SSS demonstrates some construct validity. However, the initial 18-item instrument loaded onto two separate factors, which lowers some of the confidence in its construct validity. The two factors were split between positively phrased and negatively phrased items on the FIDES instrument. This indicates that the negatively phrased items were interpreted by participants differently than the positively phrased items, and may indicate that the negative items are not tapping into the same underlying psychological construct of their positively phrased counterparts. Psychometrically, a survey that intends to measure interest as one single construct will need to have its items load on one single factor in order to support its construct validity and the negatively phrased items on the initial version of the FIDES survey do not demonstrate this construct validity. Thus, the FIDES 1.0 scale scores were computed using only the 12 positively phrased items. However, from a pragmatic perspective, negatively phrased items were useful in identifying irregular response patterns that could indicate lack of engagement or careless reporting.

The FIDES 1.0 instrument scores were not distributed in a way that is supported by interest theory. After removing the negatively phrased items, the FIDES 1.0 instrument had a skew (i.e., the distribution of scores is considerably lower on one side of the peak than on the

other) slightly towards the *strongly disagree* side of the scale, which would indicate that responders were more likely to be on the higher end of the scale. In general, a survey based on the FPMID would be expected to skew towards the *strongly agree* side of the scale since very high levels of (well-developed) interest in engineering are far more rare in the population as a whole (particularly a sample of historically under-represented students) than low (triggered situational) interest levels[1]. A survey that is sensitive to all of the FPMID's four phases of interest will need to better model this distribution, which includes a mean below the center point of the scale and a skew to the *strongly agree* side of the scale. Without such a skew, a ceiling effect on FIDES scores may result for responders who are at high levels of interest.

While most of the group comparison results behaved as expected – with no group differences for teacher or grade level, and significant difference between genders - results from comparing group means on the FIDES 1.0 lend further support for the need for revisions the FIDES 1.0. The major flaw found during our group comparisons, was that we did not find a significant difference between students who intent to major in STEM fields and their peers who plan a major outside of STEM. A survey that intends to measure student interest in engineering as a trait, should be able to distinguish students indicating future interest in STEM from those who do not. This finding indicates that a need to refine the FIDES 1.0 in order to measure interest in engineering as a psychological construct in a way that more accurately reflects our understanding of the intended population.

FIDES 2.0: Revised Instrument Development

Revised Item Construction

Revisions were made to the FIDES instrument on the basis of results from the pilot study. First, two additional indicators were added (*content knowledge* and *self-efficacy*). Second, to address issues with the scale score distribution, definitions for previous indicators were revised to include more extreme and specific phrasing of items. Consequently, Likert-scale statements were reworded with modified phrasing to reflect new indicator definitions. As a further attempt to address the need for the scale to be distributed more towards the *strongly disagree* end, one item for each indicator had conditional phrasing removed.

A review of the FPMID, and consultation with one of the FPMID authors (K.A. Renninger, personal communication, August 5th, 2014), suggested that two additional indicators be included in the FIDES survey to more accurately measure interest based on the FPMID theory. The first indicator, *content knowledge* (KNOW) had previously been identified during the initial phase of item construction, but was not included. However, it was determined that the importance of including content knowledge as an indicator outweighed potential issues with it being measured via self-report. A second additional indicator, *self-efficacy* (SE), was also suggested as an addition to the survey. According to the FPMID self-efficacy has a reciprocal relationship with interest[7, 22], and higher levels of self-efficacy are indicative of higher phases of interest and vice versa[23]. Both new indicators were added, for a new total of seven FIDES interest indicators. Items for the KNOW indicator were developed from FPMID described learner characteristics, and items for the SE indicator were modeled after items from the Self-Efficacy portion of the Motivated Strategies for Learning Questionnaire[24].

Revisions to the definitions of the existing FIDES indicators included adding more extreme and specific wording. Phrasing emphasizing frequency was added for *content*

engagement and *independent questioning* indicators. The indicator *positive feelings* was changed to *personal value* (VAL) and includes phrasing so that the indicator implies the value of engineering is more about personal value rather than a general value to society. The definition for *use of feedback* was modified to more specifically define the type of feedback intended. Finally, the definition for *perseverance* was modified to more specifically define the conditions in which students feel they demonstrate perseverance in working on engineering projects. The final list of seven interest indicators and their definitions are listed in Table 6.

Table 6:
Final FIDES interest measurement indicator definitions.

Category	Definition
Independent content engagement (CONTENT)	The frequency with which a student *independently* (re)engages in activities within the content domain. Activities can include attending after school clubs, attending museums, readings or other pursuits of content knowledge.
Independent questioning (QUES)	The frequency with which a student *independently* develops problems or questions to be solved within the content domain. This can include pursuing study of a self-created content question, or project-type work in order to solve a self-identified domain problem e.g. creating a computer program to control a small robot.
Personal value (VAL)	The amount of value or usefulness *to themselves* that a student perceives the domain to have.
Use of feedback (FEED)	The extent to which a student prefers to find out how to answer problems or questions in the domain on their own or would rather be given the answer or clear steps on how to find a correct answer.
Perseverance (PERSEVERE)	The extent that a student is willing to continue working on a problem or project in the domain when the problem becomes very difficult or when it has taken a long time to complete.
Content knowledge (KNOW)	Self-reported measure of a student's content knowledge level compared to their peers.
Self-efficacy (SE)	A self-reported measure of a student's confidence in their ability to learn in the content domain[25].

FIDES 1.0 items included a conditional phrase which generally began each question with "when I am interested in an engineering topic…" Since the FIDES instrument is intended to measure interest for students at all phases, this conditional phrase was removed for some items, as it was deemed more appropriate for measuring student interest for those who have only situational interest. Item statements were also modified to include more specific or extreme statements in order to reduce some of the ceiling effect for higher interest respondents. For example, the FIDES 1.0 instrument included the statement:

CE3: When I find something to be interesting in engineering, I love to learn more about it when I'm outside of class.

This statement was modified for the FIDES 2.0 instrument to read:

CE1: I work on engineering projects outside of school at least once a week.

The revised item removes the conditional phrase, "When I find something to be interesting in engineering…", and is more specifically worded by stating a definite time frame for work outside of class to be "at least once a week." This provides an item better suited to measuring interest for students who are at a higher (individual) phase of interest. The FIDES 2.0 items reflect a better balance than FIDES 1.0 of questions that apply to both situational (triggered and maintained) and individual (emerging and well-developed) levels of student interest. See Table 7 for a complete list of FIDES 2.0 items.

Table 7: FIDES 2.0 Instrument Items

Item Label	Statement
PERSEVERE 1	I enjoy learning about engineering even when it is very difficult.
PERSEVERE 2	When I'm working on something in engineering that I think is interesting, I continue working even when it takes a lot of time.
CONTENT 1	I work on engineering projects outside of school at least once a week.
CONTENT 2	I always learn more about engineering on my own if I find it interesting
VALUE 1	Knowing about engineering is extremely valuable to me.
VALUE 2	I think everyone should know a lot about engineering
QUES 1	I think of my own engineering projects at least once a week.
QUES 2	I'm inspired to come up with my own engineering projects to work on when I see something in engineering that interests me.
FEED 1*	When I'm working on an engineering project, I prefer to be told how to do the work.
FEED 2*	When I'm working on an engineering project that I find interesting, I like for teachers to show me what to do.
KNOW 1	I know way more about engineering than other kids I know.
KNOW 2	I know a lot about the engineering topics that I find interesting.
SE 1	Compared to other students at my school, I am way better at doing engineering work.
SE 2	When engineering interests me, I am confident that I can learn about it extremely easily.

* Items removed from final FIDES instrument.

Revised Survey Administration

The revised survey consisting of the 14-item FIDES instrument, seven negatively phrased irregular response indicators, seven E-SSS items for ascertaining construct validity, and four demographic items (current science teacher information was not assessed) was administered to 145 students (48 male, 97 female). Initially, testing the revised FIDES instrument was to be done using a sample from the original participating school (from Study 1) as well as an from an additional school. Due to weather related school closings the original school was not able to participate in Study 2. As a result, the revised FIDES (FIDES 2.0) was tested using a sample

from the additional school. The school is a small urban charter high school in a large Midwestern city serving a predominantly Latino/a population (98%), with a high rate of students receiving free or reduced lunches (84%). The school has 85% of its student body meet or exceed expectations on the state standardized test. Students represented all high school grade levels (9^{th} = 33, 10^{th} = 41, 11^{th} = 34, 12^{th} = 37). The survey was administered in one online session conducted at the school site. All of the responses were used in the subsequent analysis since no set of responses reached the pre-established criteria for irregular response patterns in which a participant's response to the positively phrased items was the same as their response for the negatively phrased item for each indicator.

Validity and Reliability Analyses – FIDES 2.0

As in the pilot study, a principal components factor analysis was performed to assess whether all 14 FIDES items loaded on a single factor. Again, similar to the pilot study, the initial analysis showed that one factor explained 42.7% percent of the variance, but a second factor explained an additional 12.0% of variance (see Table 8). A two-factor analysis was then conducted using all 14 FIDES items. All FIDES 2.0 items loaded at a practically significant level of 0.5 or higher within factor 1 (range of 0.543 to 0.787), except for the two items used to measure the indicator *use of feedback*, FEED1 and FEED2, which had loadings of 0.088 and 0.281, respectively. The two items both loaded significantly within factor 2, with loadings of 0.753 to 0.748, respectively. See Table 9 for a full list of factor loadings.

Table 8:
Factor analysis of FIDES 2.0 survey, variance explained.

Component	Variance Explained
1	42.7
2	12.0
3	7.6

Table 9:
Factor analysis loadings for FIDES 2.0 items.

Item	Factor	
	1	2
PERSEVERE 1	.662	.122
PERSEVERE 2	.664	.356
CONTENT 1	.706	-.280
CONTENT 2	.687	-.041
VALUE 1	.737	-.003
VALUE 2	.543	.113
QUES 1	.708	-.083
QUES 2	.787	.054
FEED 1	.088	.753
FEED2	.281	.748
KNOW 1	.692	-.353
KNOW 2	.774	-.190
SE 1	.761	-.265
SE 2	.654	.280

Since the factor analysis indicated that the 14 item FIDES 2.0 survey loaded on two separate and mostly distinct factors, only the 12 items that loaded onto factor 1 were used in calculating FIDES 2.0 scale scores. The scale demonstrated high reliability ($\alpha = .89$). Scale scores were centered left of the middle point of the scale ($\bar{X} = 3.23$, $\sigma = 1.21$; see table 10), and represented a wide range of scores (1.0 to 7.0). The distribution of the scores was modeled by a normal distribution ($W = 0.983$, $p = 0.07$; see Figure 2) with a large skew to the *strongly agree* side of the scale (0.434) and a Kurtosis (-0.045) close to Gaussian shape. FIDES scores were also highly correlated with the E-SSS results ($r = 0.69$). Means for each of the six remaining interest indicators were all within 0.5 of the overall FIDES 2.0 mean score (see table 10).

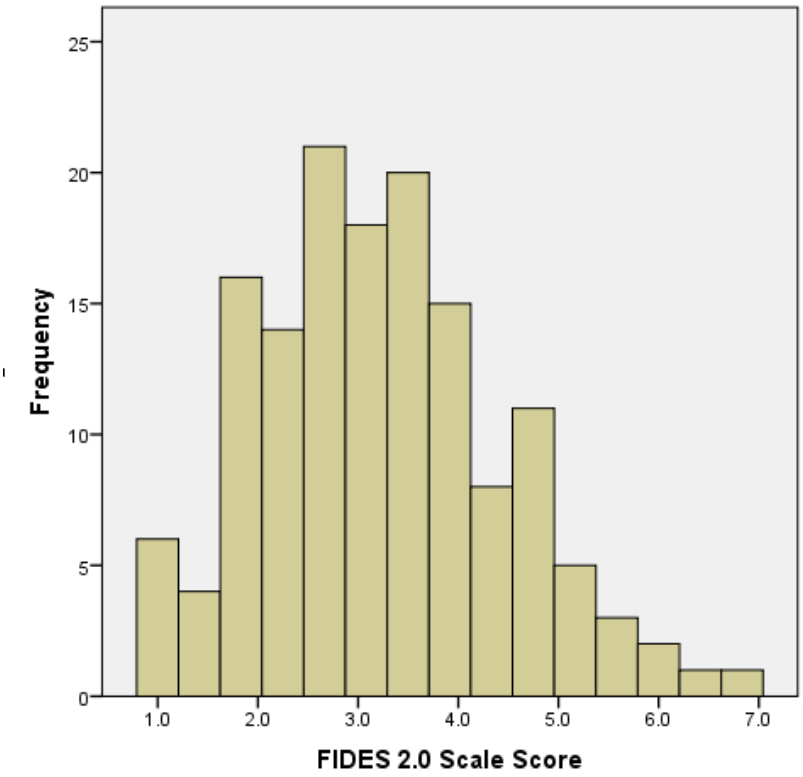

Figure 2: FIDES 2.0 scale score distribution.

Table 10: Descriptive statistics for FIDES 2.0 Indicators

Indictor	# of Items	Mean	σ	α
PERSEVERANCE	2	3.89	1.54	0.63
CONTENT	2	2.76	1.45	0.62
VAL	2	3.76	1.46	0.68
QUES	2	2.77	1.53	0.69
FEED*	2	4.20	1.49	0.64
KNOW	2	2.93	1.51	0.72
SE	2	3.32	1.43	0.59
FIDES 2.0	12	3.23	1.21	0.91

*Note: The *use of feedback* indicator was not included in the final FIDES 2.0 scale.

Revised Results

Again, after having established the reliability and the validity of the FIDES 2.0 instrument, we examined how reported interest varied by grade level, sex, and intended academic major, based on FIDES 2.0 scale scores using the 12 remaining items (after having removed the FEED1 and FEED2 items). See Table 11 for scale score means by group.

Table 11:
Descriptive statistics for mean FIDES 2.0 scores, by demographic grouping

Group	Count	Average	σ
Grade			
9	33	3.3	1.26
10	41	3.55	1.16
11	34	3.21	1.32
12	37	2.83	1.04
Gender			
Male	48	3.68	1.32
Female	97	3.01	1.11
Major			
Undecided	73	3.12	1.13
STEM	35	3.71	1.38
Humanities	21	2.8	1.18
Business	12	3.18	1.26
Performing Arts	4	3.41	0.40

As with the pilot data set, there were no significant differences in mean FIDES 2.0 scores based on grade level (F (3,140) = 2.36, p = 0.74), or intended major (F (4,139) = 2.32, p = 0.60). However, in contrast to the pilot data, a comparison of the FIDES 2.0 mean for students intending on majoring in a STEM field (3.72) was contrasted with all other majors (2.96; not including undecided) and a significant difference was found (t(71) = 2.55, $p < 0.05$). As the mean scale scores indicate, students who intend to pursue STEM majors report a higher level of interest than those who do not intend to declare a STEM major. Although this is not surprising, it provides further support for the construct validity of the FIDES 2.0 instrument.

Mean FIDES 2.0 scores for males (3.68) and females (3.01) were found to be significantly different (t (144) = 3.16, $p < 0.05$). The mean scale scores can be interpreted as suggesting that males tend to occupy a higher interest level than female students in this sample.

Discussion of Revised Item Testing

The FIDES 2.0 instrument performed as well as the FIDES 1.0 instrument in both reliability and in construct validity. The high Cronbach's alpha ($\alpha = .89$) showed a continued strong internal reliability and indicated close correlation between items on the scale. The high correlation (r = 0.69) with the independent measure of student interest in engineering, the E-SSS, also provides empirical evidence for construct validity. Further support for construct validity exists in evidence from the factor analysis performed on the 12 item FIDES 2.0 instrument, after analyzing using two factors.

The major issue addressed with revisions of the FIDES 1.0 instrument was to change the shape of the scale score distribution. As mentioned in the discussion of the results for the FIDES 1.0, interest in engineering is expected to be low for a general high school population. The initial scale scores were too close to the center value and skewed slightly towards the *strongly disagree* side of the scale. We determined that the scale should have a much higher skew to the *strongly*

agree side of the scale, and have a mean closer to the low end of the interest scale. This was done in order to reduce the potential for a ceiling effect for higher interest students and to more accurately reflect the frequency with which historically underrepresented high school students show high interest in engineering. This goal was achieved through the FIDES revisions by introducing more specific and extreme language that included introducing items aimed at both low and high interest level students.

Further evidence that scale responses model the intended population exists in considering group comparisons for the scale. As with the FIDES 1.0 results, we found a significant difference between male and female respondents, but not for grade level. These results are consistent with the intended population. Also, the FIDES 2.0 testing showed a significant difference between students intending to major in a STEM field and those who are planning a major outside of STEM. This was not the case for FIDES 1.0, and is more consistent with how we would expect the population to be distributed.

The improved version of the FIDES instrument demonstrates both validity and reliability as a way to measure interest in engineering as a construct. The FIDES 2.0 survey scores are distributed in a way that allows administrators of the FIDES instrument to better measure growth of student interest over time.

Significance

This study reports on the iterative development of a way to practically and efficiently measure interest development in engineering. The FIDES 2.0 survey is recommended for use in any engineering education setting in which obtaining a solitary estimate of the current interest level of a student is of value, or which intends to evaluate program impact on changes in student interest in engineering. The survey can be used as a standalone measurement tool in order to provide programs and educators information about the current level of interest that an individual student is at. This information can be used to provide differentiated teaching, scaffolding, or grouping for individual students to best meet their developmental needs based on their interest level. Using the FIDES 2.0 as a pre- and post-intervention measure of student interest can also provide evidence used to evaluate program effectiveness in increasing student interest in engineering over time. The FIDES survey is particularly useful in this application, as it is sensitive to changes in multiple indicators of interest, and has a low risk of a ceiling effect impacting measuring change in students who began at an already elevated level of interest.

Conclusions

In conclusion, our goal for developing a tool that provides a standardized and efficient measure for interest in engineering that is based on current interest development theory, was achieved through the iterative development of the FIDES survey instrument. FIDES 2.0 is sensitive to student interest as both a state and a trait in that it utilizes items aimed at both situational and individual phases of interest described in the FPMID. To measure at all phases, the survey includes indicators of interest beyond simple positive feelings or attitudes, which the FPMID describes as only one of the characteristics that provides evidence for the level of student interest development for a domain. Examples of the broader utilization of multiple interest characteristics are items on the FIDES 2.0 based on indicators such as content knowledge and independent questioning that help evaluate student interest at higher phases. Overall, the FIDES

2.0 survey's consistency with current theory improves its ability to accurately measure interest in engineering as a psychological construct.

Beyond being theoretically consistent, the survey also demonstrates both validity and reliability, and via examination of a representative sample, demonstrates responses believed to be consistent with our understanding of a high school student population. An accurate picture of interest in engineering among historically under-represented, urban high school students includes a lower percentage of students at high levels of interest. Revision to the FIDES 1.0 created a survey instrument (FIDES 2.0) that reflects this expectation by demonstrating scale scores that result in a mean that maps towards the *slightly disagree* side of the interest scale. Furthermore, the revisions to the FIDES resulted in a distribution of scale scores consistent with our understanding of the population, as the *strongly disagree* side of the FIDES 2.0 scale scores has a precipitous drop in frequency as scores approach very high levels.

There are, however, a few limitations of the study. To improve confidence in the survey's overall generalizability, larger and more diverse samples should be tested. The study administered the survey in high schools that serve predominately minority students, and are in large urban areas. The well-documented underrepresentation of minorities in STEM careers may imply that interest in engineering fields is lower for these groups, and may have an effect on response patterns for the students surveyed. Therefore, studies that include a better balance of ethnicities are necessary. There may also be some influence of living in a large urban area on response patterns, and so testing in suburban and rural areas is also needed to fully evaluate the FIDES 2.0 instrument. Thus far, the survey appears to reflect the response pattern that would be expected for the high school student population as a whole, but the recommended further testing will strengthen the validity of that claim.

Future studies in which truncated versions of the survey have been administered are also needed to help better achieve the goal of providing an efficient measurement tool. While we believe the survey is well balanced and fairly brief, it may be necessary in some situations to have a tool that can be used very quickly. For example, in a learning task that includes a long session with an online intelligent tutor, reducing the amount of time spent on other extraneous online items may be desirable. Studies that analyze the items on the FIDES 2.0 survey to identify those that provide the most information about each respondent's scale score could reveal a smaller subset of items that would give survey administrators a shortened version of the FIDES instrument, thus giving more flexibility in its use. In its current formulation, the FIDES 2.0 instrument can be used as a stand-alone 12 item survey, with 6 optional negatively phrased items available to identify irregular patterns. The 12 items provide a brief survey, however, we recommend the addition of the negatively phrased items, as they have proven useful in detecting irregular patterns.

We believe that we have developed and refined a valid and reliable survey instrument that can provide critical information about student interest in engineering to both researchers and practitioners in engineering education.

References

1. National Academy of Sciences (2007). *Rising above the gathering storm: Energizing and employing America for a brighter future*. Washington, DC: National Academies Press.

2. National Research Council Committee on Integrated STEM Education (2014). *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. Margaret Honey, Greg Pearson, and Heidi Schweingruber (Eds.). Washington, DC: The National Academies Press.
3. Tyler-Wood, T., Knezek, G. & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education*, 18(2), 341-363.
4. Faber, M., Unfried, A., Weibe, E.N., Corn, J., Walker Townsend, L., &Collins, T.L., (2013, June). *Student attitudes toward STEM: The development of upper elementary school and middle/high school student surveys*. Paper presented at the ASEE Conference & Exposition, Atlanta, GA.
5. Maltese, A. V., Melki, C. S. and Wiebke, H. L. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education*, 98(6), 937–962.
6. Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231.
7. Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41, 111–1270.
8. Reigeluth, C. M. (1983). *Instructional-design theories and models: An overview of their current status.* Hillsdale, NJ: Lawrence Erlbaum Associates, pg. 32.
9. Ainley, M., Hillman, K., & Hidi, S. (2002). Gender and interest processes in response to literary texts: situational and individual interest. *Learning and Instruction*. 12(4), 411-428.
10. Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383-409.
11. Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13(1), 23-52.
12. Perry, B. D., Pollard, R. A., Blakley, T. L., Baker, W. L., & Vigilante, D. (1995). Childhood trauma, the neurobiology of adaptation, and use dependent development of the brain: How "states" become "traits". *Infant mental health journal*, *16*(4), 271-291.
13. Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of educational psychology*, *92*(2), 316.
14. Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, *31*(3), 313-327.
15. Crocker, L. & Algina, J. (2009). *Introduction to classical and modern test theory*. Mason, OH: Cengage Learning.
16. Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. Psychological Assessment, 7(3), 309-319.
17. Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, 15(5), 381–395.
18. Schiefele, U. (2001). *The role of interest in motivation and learning*. In J. M. Collis, & S. Messick (Eds.), Intelligence and personality (pp. 163-194). Mahwah, NJ: Erlbaum.
19. 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.
20. Lent, R.W., Brown, S.D. (1996). Social cognitive approach to career development: An overview. *The career development quarterly*, 44(4), 310-321.
21. Hair, J.F., Tatham, R.L., Anderson, R.E. and Black, W. (1998) *Multivariate data analysis*. (Fifth Ed.) Prentice-Hall:London.
22. Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, 46(3), 168-184.
23. Renninger, K. A. (2010).*Working with and cultivating interest, self-efficacy, and self-regulation*. In D. Preiss & R. Sternberg (Eds.), Innovations in educational psychology: Perspectives on learning, teaching and human development (pp. 107–138). New York, NY: Springer.
24. Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.

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