Development and Testing of an Instrument to Understand Engineering Doctoral Students' Identities and Motivations

Heather Lee Perkins, North Carolina State University

Heather entered the Applied Social and Community Psychology program in the fall of 2014, after completing her Bachelor of Science in Psychology from the University of Cincinnati. She has participated in various research projects examining the interaction between stereotypes and science interest and confidence, their influence upon womens' performance in school and the workplace, and their presence in the media and consequences for viewers. Her primary research interest is science identity, STEM education, and participation in online communities.

Mr. Matthew Bahnson, North Carolina State University

Matthew Bahnson is a doctoral student at North Carolina State University in Applied Social and Community Psychology. His research interests include engineering identity, diversity, bias, stereotypes, and STEM education. He works with Dr. Cheryl Cass at NCSU.

Mrs. Marissa A. Tsugawa-Nieves, University of Nevada, Reno

Marissa Tsugawa is a graduate research assistant studying at the University of Nevada, Reno in the PRiDE Research Group. She is currently working towards a Ph.D. in Engineering Education. She expects to graduate May of 2019. Her research interests include student development of identity and motivation in graduate engineering environments and understanding creativity in engineering design processes.

Dr. Adam Kirn, University of Nevada, Reno

Adam Kirn is an Assistant Professor of Engineering Education at University of Nevada, Reno. His research focuses on the interactions between engineering cultures, student motivation, and their learning experiences. His projects involve the study of student perceptions, beliefs and attitudes towards becoming engineers, their problem solving processes, and cultural fit. His education includes a B.S. in Biomedical Engineering from Rose-Hulman Institute of Technology, a M.S. in Bioengineering and Ph.D. in Engineering and Science Education from Clemson University.

Dr. Cheryl Cass, North Carolina State University

Cheryl Cass is a teaching assistant professor in the Department of Materials Science and Engineering at North Carolina State University where she has served as the Director of Undergraduate Programs since 2011. Her research focuses on the intersection of science and engineering identity in post-secondary and graduate level programs.

©American Society for Engineering Education, 2018

Development and Testing of an Instrument to Understand Engineering Doctoral Students' Identities and Motivations

Abstract

This research paper explores the formation and psychometrics of a survey developed to understand the identities and motivations of engineering graduate students. Within STEM, almost 50% of students leave before completing their PhD, and many interventions suggest strengthening students' identities to prevent dropout. However, engineering identity and identity-based motivation (IBM) within the engineering graduate population is not well-explored. To remedy this, two questionnaires were created, with items derived from qualitative research with graduate students. Data were collected from 227 engineering graduate students, and were analyzed for patterns in missingness (i.e., individual items skipped by participants despite overall survey completion), effects on participants' emotional state (using the Positive and Negative Affect Schedule, PANAS), and validity and reliability. MANOVA analysis indicated there was no significant effect of survey focus on positive or negative affect, $F(4,646) = 1.075$, $p = .368$, but that data was not missing at random in the IBM survey, $\chi^2(503) = 580.80$, $p = .009$). With Exploratory Factor Analysis, the latent constructs were tested and the list of items refined. The implications of these findings for the full survey and for future studies will be discussed.

Introduction: the GRADS Project

Amidst calls for an increase in STEM graduates in the U.S., attrition among engineering graduate students remains a serious issue [1]. Previous studies have indicated that the attrition rate among STEM doctoral students is as high as 50% [2], and retention of students from traditionally marginalized groups continues to be of special concern [3]. These studies also indicate that strong engineering identities and clear future goals are critical to student success [4]–[6], but often fail to include graduate students as a population distinct from undergraduate students [7]–[9]. To begin remedying this gap, the GRADS project was proposed, a qualitative and quantitative investigation of engineering doctoral students' (EDS) experiences, identities, and motivation [10].

As the first step in this process, three qualitative studies were conducted with an EDS sample [11]–[13]. This was done both to investigate whether EDS framed their experiences and identities similarly to previous undergraduate populations, and to begin exploring questions for the planned quantitative phase. Results indicated that engineering identity, identity fit, and future goals were as important for doctoral students as they were for undergraduates, but that the constructs were sometimes articulated and prioritized differently. For instance, graduate students who strongly identified as engineers took a more agentic role in defining what 'engineering' meant, recognizing that the field is broad and that the stereotypical definitions can be limiting [11]. Similarly, graduate students used past experiences to clarify goals and describe their future place in the field, unlike undergraduates who often only considered present performance [12].

The second step in the GRADS project entailed the development and administration of a nationally representative survey of EDS. To do this, the research team gathered previously developed scales and surveys, published literature about EDS identities and motivations (including the three qualitative studies mentioned previously), and combined them to fit the current survey's goals while retaining previously established metrics of validity and reliability. This resulted in the development of three full-length surveys, each focusing on engineering identity, identity-based motivation, and future time perspective. To combine these three surveys into a single, smaller survey that could be used with a national population, we piloted them with an engineering graduate student sample and the results were analyzed for validity and reliability using exploratory factor analysis (EFA), ANOVAs, and Little's MCAR test (missing completely-at-random test).

This research paper explores the development of two of the three pilot surveys, specifically those focusing on issues of identity and identity-based motivation. The process by which previous research and interview data from graduate students were combined is explored, as well as the quantitative analyses of the surveys' psychometrics. We conclude with recommendations for others seeking to extend and transfer findings from the undergraduate to the graduate space, as well those seeking to develop surveys for use with nationally representative populations of engineering students. Overall, the goal of this paper is to recount the creation, testing, and validation of the survey measure before it was used with a national population. We hope this paper will contribute to the discussion of survey development best practices within engineering education, as well as act as a future resource for the GRADS project as data and results are shared.

Method

Measures

Survey Item Development. Two surveys were developed to explore EDS's engineering identities and how their identities impacted motivation in their doctoral programs. Items were drawn from existing research, specifically Godwin's (2016) work on engineering identity and Oyserman's (2015) theory of identity-based motivation. Because most of this work focused on undergraduate students, data gathered from qualitative interviews was used to guide the process of modifying and tweaking the items for use with a graduate population (for a full qualitative analysis of identity and identity-based motivation among EDS, see Perkins et al. (2017) and Miller et al. (2017)).

Engineering Identity Items. Previous publications have explored engineering identity among undergraduate students and identified three factors central to the construct: performance/competence, interest, and recognition [4]. In previous research, these factors were examined via a series of questions that asked participants to rate, on a seven-point Likert scale, how much they agreed with statements such as the following: "I understand concepts I have studied in engineering"; "I am interested in learning more about engineering"; and "My peers see me as an engineer". To explore the intersections between math, physics, and engineering identities, participants were also asked identical questions rephrased to explore math and physics domains [15].

When interviewing EDS about their engineering identities and graduate experiences, several trends were noticed. As was reported in Perkins et al. (2017), EDS often questioned whether they were recognized as legitimate engineers by others in the field, and theorized that they were viewed differently depending on others' roles rather than their fields (i.e., advisors and faculty and non-academic engineering professionals were expected to have differing and sometimes incompatible definitions of engineering). Several participants also expressed the view that they could be defined as 'scientists' rather than 'engineers' and expressed some doubts as to which label they preferred, both of which impacted how accepted they felt in their programs and their perceived progress. As part of these discussions, participants also invoked their identities as researchers in contrast to an ill-defined stereotype of a traditional 'professional' or 'industry engineer', indicating polarity between knowledge-building and sharing versus hands-on and applied work. In short, EDS still struggled with their engineering identities, but there was less focus on performance/competence and interest, and more focus on recognition. Similarly, there was less focus on their knowledge domains, and more focus on their values and the type of work they performed.

As a result, the previously existing items with their established psychometrics were modified slightly for use in the new population. Some items were tweaked to be more descriptive of an EDS experiences, e.g. "I can do well on exams in SCIENCE" was revised to read, "I can perform well when my SCIENCE knowledge is tested (for instance, in exams or defenses)". Additionally, rather than create sets of parallel items to ask about participants identities within the domains of engineering, science, and math, items were instead created to ask about their identities as scientists, engineers, and researchers (e.g., the item "Other engineers see me as an ENGINEER" was mirrored by "Other scientists see me as a SCIENTIST" and "Other researchers see me as a RESEARCHER").

This resulted in 61 items: 13 items paralleled to query students about their engineering, scientist, and researcher identities, and 14 items that examined specific aspects of engineering and researcher identities indicated as significant in the interview data (see Tables A1 – A3 for a full list of engineering identity items). Participants responded to each item on a 5-point Likert scale, with lower scores indicating less agreement. Although previous surveys with undergraduates used a 7-point Likert scale, the current survey limited Likert items to five response options. This was done in light of research that indicates five-item scales induce less misresponse in reversed items, an issue of particular concern for our survey given the high number of international students in engineering doctoral programs [16].

Identity-Based Motivation Items. Items for the identity-based motivation (IBM) survey were developed based on Oyserman's theory of Identity-Based Motivation (IBM) and the results from the authors' previous qualitative study [13]. Based upon Oyserman's theory, three central tenants of IBM were identified for inclusion in the survey: dynamic construction, action-readiness, and interpretation of difficulty. Dynamic construction describes the ways that identities are constructed within context, e.g., priming students to view their gender as more or less academically successful and then asking them to reflect on their future goals has been demonstrated to impact the goals listed [17]. Action-readiness is the process by which salient identities prompt individual to engage with related activities, and how these activities impact their overall motivation [14]. Interpretation of difficulty refers to the ways in which students respond to failure. In the case of identity-congruent tasks, it signals that the task and identity are important and require more effort. For identity-incongruent tasks, failure indicates that the identity is unlikely or unimportant, and one should withdraw from the task.

To assess dynamic construction among EDS, salient identities and relevant contexts were drawn from the qualitative data. Ten items asking about salient identities when performing identity-relevant tasks were developed, with three identities (scientist, researcher, and engineer; the rationale for focusing on these three identities is outlined in the section above) provided as responses and a fourth 'other' option available for students to write-in an identity that was not included (see Tables A4 and A5 for a full list of IBM items). For each option, participants rated how much they "felt" like the specified identity when performing the specified task on a 5-point Likert scale. To explore action-readiness and interpretation of difficulty, participants were provided a list of 9 identity-relevant tasks and asked to indicate how much each aligned with their overall identity and how difficult they found the tasks to be, totaling 18 items, each on a 5-point Likert scale. The final IBM survey was 48 items long, with 30 items focusing on dynamic construction and 18 items focusing on action-readiness and interpretation of difficulty.

Positive and Negative Affect Schedule. Once participants completed the bank of test items, they were asked to complete the short form of the Positive and Negative Affect Schedule (PANAS) [18], [19]. Based on the qualitative sessions and reviews of the literature, the concern was raised that participants' moods may be negatively impacted by different survey's subject matter, thus systematically influencing their responses. For instance, previous studies have indicated the importance of belonging and STEM identity to students' evaluations of their engagement and self-efficacy in the classroom, and so it was suggested that priming students to think about their engineering identity may impact their responses to items querying their degree progress or future goals [20].

To determine if the final survey should use counterbalancing to prevent earlier questions from biasing responses to later items, the PANAS was used to screen for differences in mood, either positive or negative, between students who completed the different pilot surveys. The I-PANAS-SF is a short form of instrument that has been developed and tested with an international population; it consists of ten items (comprised of five positive and five negative emotion words) and asks participants to report how much they experienced each feeling within the specified time frame on a five-point Likert scale [18].

Participants

With construction of the pilot instrument complete, 1,971 EDS at two land-grant institutions on the east and west coasts were invited to participate, with a participation rate of approximately 14%. The final tally across the two surveys analyzed here (identity, n = 107; IBM, n = 120) showed the demographics in line with those reported by national surveys [21], with the majority identifying as male (63%), Asian (42%) or White (38%), international (62%), and heterosexual (93%). The majority of participants had not completed any of their degree milestones (68%), and very few had proceeded to their dissertation defense (2%), with the remainder split between completion of their comprehensive exams and dissertation proposals (17% and 10%, respectively).

Analyses

To assess the effectiveness of the survey instrument before use with a national population, the data were analyzed for patterns in missingness, the effect of the surveys on participants' emotional states, and for adherence to previously demonstrated patterns of validity. Of the participants who reached the survey's end, only 279 responded to all items fully, resulting in 54 participants with missing data. Little's MCAR (Missing Completely At Random) test was used to determine if items were omitted randomly, or if there were systematic patterns in participant's omissions. This is done following best practices in the literature, as data missing NAR (Not At Random) can interfere with statistical analyses and so researchers are advised to analyze missingness patterns before running test statistics [22].

Once the patterns in missing data were analyzed, responses to the PANAS items were averaged into two categories, positive and negative. Since each participant completed only one of the three pilots, a 3x2 MANOVA (identity, IBM, and FTP survey x positive and negative affect) was run to test for effects of survey type on participants' emotional state (i.e., positive or negative affect).

Last, the test items were evaluated using Exploratory Factor Analysis (EFA) according to recommendations in the literature [23], [24]. Factor analysis, which includes exploratory factory analysis (EFA, used in this study) and confirmatory factor analysis (not used here), is an estimation process that models the relationships between related items. This process allows for the identification of shared variance, also known as latent variables or factors, as well as unique variance. Shared variance is the extent to which items co-vary: if items have high covariance, it is assumed that the changes in the item scores are caused by the same psychological construct. For instance, an item that evaluates whether one's peers see one as a scientist, and a separate item that evaluates whether one's advisor sees one as a scientist, have high covariance as they are both impacted by the underlying construct of 'recognition'.

A third item that evaluates one's confidence in a science class does not have high covariance with the first two items, as it is impacted by the underlying construct of 'performance/competence'. A fourth item that evaluates one's understanding of science concepts would co-vary greatly with the third item, but not with the first two; in these patterns the underlying factors, or constructs, become visible. (Unique variance, mentioned above, is the 'leftover' variance that cannot be explained by the latent variable; e.g., responses to the second group of questions would be influenced by 'performance/competence' in science but also 'performance/competence' in other areas, such as the English language or in academic settings.)

Although EFA is not a confirmatory test – that is, it does not test a hypothesis and produce fit statistics – it is still used to test and confirm assumptions about measures and items. In this study, we drew from previous research with undergraduates and interviews with graduate students to develop items that explore engineering identity and identity-based motivation (IBM) among EDS. For the engineering identity items, we had predictions about how they would behave, e.g., that they would load on established factors of recognition, performance/competence, and interest.

For the IBM items, we developed the scales specifically for this survey, and while we did so with the assumption that items would load onto specific patterns, we did not have clearly defined predictions. To 'test' these predictions and assumptions, we ran a series of EFAs to explore the underlying patterns formed by participants' responses and to guide the process of shortening and streamlining the overall survey.

Although participants completed all items in both surveys simultaneously, the items were designed to operate as independent measures of related constructs and so were tested in five separate EFAs (e.g., the three constructs of engineering identity). This not only allowed for clearer interpretation of results, but ensured that there was a sufficient participant to item ratio [25], [26]. The maximum likelihood estimation method with Promax rotation was used for all EFA work, allowing for correlations between factors (as all underlying constructs were theorized to be related to one another). Resulting scree tests, parallel analyses, and factor loadings were used to assess the final solutions [27]. The primary focus of this analysis was determining whether established constructs (e.g., competence/performance and action-readiness) behaved as predicted within a new population. These tests worked both to demonstrate the validity of the underlying constructs and the reliability of the survey instrument.

Results

Missingness

Before analyzing the data for patterns in missingness, a visual review was conducted to look for patterns in skipped questions. As 27% of the participants skipped one or more items asking about teaching assistantships, all teaching items were removed from analysis, resulting in a new total of 43 items for the IBM survey (the identity survey was unaffected). The remaining items in both surveys were completed by at least 90% of the sample, and so were retained.

Little's MCAR test indicated that unanswered items were missing completely at random for all engineering identity items (p = .17) but not in the identity-based motivation survey, $\chi^2(503, n = 120) = 580.80$, p = .009). To determine whether this pattern reflected biases along demographics of interest, regressions were used to test missingness patterns across the three primary demographic categories (race, gender, and international status). Results indicated that men, international students, and Asian students were significantly more likely to skip at least one of four items: three items which asked about their salient identities when writing peer-reviewed papers, and one that asked about participants' overall identity as a scientist (see Table A6 for item text and regression values).

Positive and Negative Affect

A 3x2 MANOVA (survey type x affect type) was run to test for effects of survey type on participants' emotional state (i.e., positive or negative affect). Results indicated there was no significant effect of survey focus (i.e., engineering identity, identity-based motivation, and future time perspective) on positive or negative affect, $F(4,646) = 1.075$, p = .368 (see Table 1 for full means, standard deviations, and univariate effects). These results indicate that there is not a need for counterbalancing, as emotional state was not impacted by survey content.

Table 1. Effects of Survey Topic on Participants' Positive and Negative Affect

Affect Type	Survey Topic	*M*	*SD*	*df*	*F*	*p*
Negative	Identity	1.892	0.09	2, 323	1.119	0.328
	Identity-Based Motivation	2.048	0.087			
	Future Time Perspective	1.93	0.092			
Positive	Identity	3.398	0.099	2, 323	0.855	0.005
	Identity-Based Motivation	3.3	0.095			
	Future Time Perspective	3.186	0.101			

Note: MANOVA was not significant at the multivariate level, $F(4,646) = 1.075$, $p = .368$.

Exploratory Factor Analysis

Engineering Identity. The full list of 61 items was divided into three categories, focusing on domains of scientist identity, engineer identity, and researcher identity, respectively. Previous analyses of the engineering identity domain with undergraduates revealed a three-factor solution, in which items formed constructs of recognition, competence/performance, and interest [4]. Scree plot analysis of the scientist and engineer blocks indicated that a three-factor solution was adequate, as predicted. Analysis of the researcher block suggested that a two-factor solution might be adequate, although the number of eigenvalues greater than 1 was three, in line with previous solutions.

The scientist items loaded as predicted, with only one item near the.32 threshold recommended in the literature [23] (see Table A1). The engineer block was not as clear, with three Heywood cases (items with factor loadings greater than 1.0, a mathematical impossibility that suggests extreme trouble with the solution, [23]) presenting in the initial solution. Two of the three items were dropped, producing the results reported in Table A2, with most items loading on their established factors. Finally, two- and three-factor solutions for the researcher block were attempted, with the three-factor solution producing the best fit. Most items loaded according to established constructs, with three items showing some problematic cross-loading. Of these items, two loaded primarily as predicted, with cross-loading occurring near the .32 threshold, and the third item loaded equally on two factors (see Table A3).

Identity-Based Motivation. The full list of 43 items was divided into two categories, one for items that asked about salient identities and one for items that asked about task alignment and difficulty. For the first category, a review of the scree plot and parallel analysis suggested that anywhere from three to five factors could be contained within the items. After testing all three solutions, the results suggested that a four-factor solution was best (see Tables A4 and A5 for factor loadings). Most of the items loaded in factors characterized by the salient identity (e.g., scientist, engineer) except for five items that asked about schoolwork (e.g., attending class and doing homework). These items loaded together on a single factor, regardless of the salient identity specified.

The second category of items (focusing on task alignment and difficulty) loaded less cleanly. The scree plot and parallel analysis recommended four factors, and all but one item loaded cleanly on a single factor, but the factors themselves were harder to define. The alignment items (e.g., ‘The following tasks align with how I view myself: Presenting my research’) produced a pattern that suggested participants differentiated between research tasks and student tasks (see Table A5). The difficulty items (e.g., ‘I consider the following tasks to be difficult: Writing about my research’) were sorted into communication-oriented and production-oriented tasks. In addition to analyzing the factor loadings, the factor correlations (i.e., the extent to which the latent variables or factors correlate with one another) were also inspected. The alignment factors did not correlate with one another, while the difficulty factors were correlated with a small effect size (see Table 2 for the full values). Additionally, the relationships between the alignment and difficulty factors were not clear.

Table 2. Factor Correlation for Identity-Based Motivation Items (Task Alignment and Difficulty)

	Aligned Research Tasks	Aligned Student Tasks	Difficult Communication Tasks	Difficult Production Tasks
Aligned Research Tasks	1.000	0.040	0.215	-0.052
Aligned Student Tasks		1.000	-0.155	-0.100
Difficult Communication Tasks			1.000	0.393
Difficult Production Tasks				1.000

Discussion

Missingness

Results of the missingness analysis indicated that there were four potentially problematic items, all within the identity-based motivation survey. The first two, 'When I write peer-reviewed papers, I feel like an engineer' and 'When I write peer-reviewed papers, I feel like a researcher' were skipped significantly more often by men. It is difficult to determine precisely what this pattern means without follow-up interviews or questions, but by looking at the literature we can generate potential hypotheses. The first is that men and boys typically perform slightly worse on measures of verbal skill when compared to women and girls [28], thereby resulting in negative attitudes (e.g., shame, anxiety, contempt) towards writing tasks that lead them to skip or disregard these items. The second is that writing skills are not typically perceived as part of an engineers’ toolkit, which instead focuses on skills in building, designing, and inventing [29], and so male participants (who traditionally identify more with the stereotypical scientist or engineer [30]) considered the item irrelevant and decided to skip it.

Similar hypotheses can be generated to explain the patterns in missingness for the third item, 'When I write peer-reviewed papers, I feel like a scientist', which was skipped more often by men and international students. International students were not more likely to skip the two previously listed items, but their results could be considered as trending significant (p = .079) or indicative of an unexplored interaction (e.g., this findings may be powered primarily by international men). This adds an additional potential hypotheses, in which English-language skills and self-efficacy may be leading male international students to skip these items more often.

The final item that was skipped did not ask about writing: instead, it asked participants to rate their agreement with the statement, 'Overall, I identify as a scientist', and was skipped more often by students who self-identified as Asian. Again, without collecting further data, it is impossible to determine why this item was skipped more often, but hypotheses can be generated. Existing model-minority stereotypes of Asian students often pigeonhole them as being adept in science (and unskilled in other areas; [31], and thus Asian students may be reluctant to identify as scientists and thus conform to the stereotype. Additionally, there exists a debate in engineering about the definitions of science vs. engineering and who 'counts' as an engineer [11]. EDS are sensitive to this debate and to the possibility of being excluded from the field for not conforming to group norms, and so this may interact with the above-mentioned stereotype to cause dissonance in Asian students that leads them to skip the item.

Without further data, we cannot reach any conclusions as to why these items were skipped, but we can outline how we have modified the survey to account for these patterns in missingness and how we will analyze these patterns in the future. During pilot testing, students were free to skip any items they chose; in response to these patterns, we implemented an additional setting in the survey which prompted students to respond to any items left blank. Students could still proceed with incomplete data, but we hoped this would remedy some of the missingness. Once data is collected from the full sample, missingness patterns can be tested again with more sophisticated techniques (such as multiple imputation or tests of measurement invariance; Brown, 2015) to determine the effects of these patterns on results and their potential meaningfulness. Additionally, a third stage of the GRADS project is devoted to following-up with participants and conducting interviews to resolve unanswered questions, i.e., how participants felt when responding to questions, how they interpreted them, and if skipped, what motivated them to do so. Although these results are potentially alarming and have negative implications for the reliability of the survey, at this point there isn't enough data to determine whether the patterns are significant enough to indicate differential functioning or test bias across groups. Instead, these results are used as a reminder and signifier of potential concerns, and will be followed-up upon as data collection and analysis proceeds.

Exploratory Factor Analysis: Engineering Identity Discussion

Scientist Identity. Items behaved as predicted on this measure, loading on the correct factors and above the recommended cut-off value. However, some items still needed to be cut, as the goal of piloting these two surveys was to shorten them and combine them in to one. As a result, we looked back to interview data gathered previously [11] and determined that the best items to cut were 'My family sees me as a SCIENTIST' and 'Others ask me for help with SCIENCE'. The first was cut due to statements from students that, as they became increasingly specialized in their graduate careers, their families were less familiar with the work they were doing and they looked to them less often for recognition and approval. The second was cut to maximize parsimony, as it potentially conflated 'recognition' with 'performance/competence' and there was already an adequate number of 'recognition' items.

Engineer Identity. Some items behaved contrary to predictions on this measure. In particular, two items ('My department faculty see me as an ENGINEER' and 'I am interested in learning more about ENGINEERING') presented as Heywood cases [23] with factor loadings above 1.0, which should not be able to occur. These items were dropped from the analysis and it was re-run, producing acceptable values for the remaining items. However, there were still issues with cross-loading. 'I can overcome setbacks when doing ENGINEERING', which was expected to load on 'performance/competence', also loaded on 'interest', and thus was dropped. Another cross-loading item, 'I see myself as an ENGINEER', also cross-loaded, but this was partially to be expected given the nature of the item (e.g., it was expected to function as an overall assessment of engineer identity, thus informed by all three constructs). As a result, this item was retained.

Researcher Identity. This measure had the most 'new' items, or items that were inspired by previous questions used with undergraduates and expected to load on the planned constructs, but modified extensively to refer to graduate students experiences. As a result, this section saw the most dropped items, as review indicated overlap between or poor functioning of planned items. Two items, 'My family sees me as a RESEARCHER' and ‘Others ask me for help with RESEARCH' mirrored earlier scientist items that were also dropped; in order to preserve similarity across the three scales, these items were discarded as well. Two items showed problematic cross-loadings, specifically 'I am confident that I can discuss current research trends' and 'I can overcome setbacks when doing RESEARCH', and were dropped. Finally, there were a number of new items created for 'performance/competence' based on previous qualitative work with graduate students [11], and given the length constraints on the survey some of them needed to be dropped. Although the following items performed adequately in the EFA, they were discarded to maximize parsimony and streamline the scale:

- I can obtain my desired RESEARCH results.
- I am confident that I can find and understand literature in my field.
- I can apply current industry trends and practices in my RESEARCH.
- I understand how to use the necessary tools/equipment for my research.
- I am confident that I can work with other researchers.
- I am confident in my ability to mentor junior members of my research group.

Exploratory Factor Analysis: Identity-Based Motivation Discussion

Salient Identities. Although the scree plot output and parallel analysis produced conflicting results, a stable four-factor solution was identified for these items. Most of the items loaded on factors related to the salient identity being explored, e.g., 'When I read journal articles, I feel like a scientist' and 'When I conduct research, I feel like a scientist' loaded together on the 'scientist' factor. However, any items that asked about school-related tasks, i.e., attending class and doing homework, loaded on a separate factor regardless of the salient identity specified. The only exception to this pattern was the item 'When I attend class, I feel like an engineer', which cross-loaded on the 'engineer' factor and the 'student' factor.

These findings challenged our previous assumptions about how the items would behave but are still congruent with theory. We initially predicted that items would load into factors defined by the type of task, e.g. items that asked about reading journal articles and writing peer-reviewed papers would load together, regardless of the salient identity specified. This aligned with the principle of dynamic construction, e.g., that students' salient identities are dynamically activated by the setting or type of task performed. However, we found that it was the salient identities themselves that emerged as the latent variables, not the types of tasks, except for student tasks which loaded on their own factor.

There are two potential interpretations of these findings. The first is that these results indicate issues with the measurement instrument, specifically that it is not accurately assessing participants' experiences of dynamic construction. The second is that EDS experience dynamic construction differently than other students. One hypothesis is that, as graduate students define themselves as experts in their content areas, their identities are less subject to interference from environment or task. This hypothesis is supported by the fourth 'student' factor, which indicated that in student settings a specific student identity is consistently activated and made salient – or, in other words, a student setting makes an EDS's student identity salient to the exclusion of their primary identity, but in other settings, they experience less dynamic construction and thus more stable 'scientist', 'engineer', or 'researcher' identities.

If this is the case, we would expect to see an absence of measurement invariance when comparing students at dissimilar stages of degree progress. For instance, we might see the initially predicted factor structure for newer students, such as undergraduate freshmen, and the currently demonstrated factor structure with more advanced students, such as those in our population. This can be evaluated using confirmatory factor analysis (CFA) with multiple groups [32], but our current data is not sufficient to test this hypothesis, and so it remains a possibility for future studies.

For the final iteration of the survey, all of the items asking about salient identities were kept, as there was little overlap between the tasks described and they largely behaved as predicted. The 'student' items were also kept, in the hope that the response patterns around these items could be further explored in the full data set. Despite the indication that 'student' is a salient identity for our participants, we decided against including a student option separate of the scientist, engineer, and researcher questions. This was done due to concern that including a student option may prompt participants to respond in accordance with their titles (e.g., "I'm a graduate student, so I should pick student") instead of engaging thoughtfully with the question. However, participants do have the option to provide low ratings to all three identities, thus providing an equivalent to a 'none of the above' option, thus allowing for a 'student' option to emerge as needed.

Task Alignment and Difficulty. Items asking about task alignment (e.g., "The following tasks align with how I view myself") loaded together, as did the items asking about task difficulty (e.g., "I consider the following tasks to be difficult"). However, each set of alignment/difficulty items loaded across two factors. Although the tasks listed in each section were the same, with only the item stem varying, the two sets of factors did not mirror each other; as was mentioned previously, the alignment items loaded into 'research task' and 'student task' factors, while the difficulty items loaded onto 'communication-oriented' and 'production-oriented' factors.

The reasons for this clustering of tasks is not clear, and is something that will be followed-up on in post-survey interview sessions. Consistent with the ideas discussed in the 'salient identity' section above, it may be that EDS do not see student tasks as consistent with their views of themselves, as they are crafting identities as experts and specialists rather than students. However, if this were the case, we would expect to see similar clustering when the tasks were framed in terms of their difficulty, e.g., 'The following tasks align with how I view myself: attending classes' and 'I consider the following tasks to be difficult: attending classes' would load on the same factor, indicating that identity-congruence and identity-incongruence were the latent variables emerging from the analysis.

The latent variables of ‘communication’ and ‘production’ tasks within the difficulty items are also unexpected. The inclusion of items that asks about disseminating research (e.g., writing and presenting) as opposed to other research tasks (e.g., collaborating and conducting research) was based on interview data, which suggested that EDS did not always feel adequately prepared for these tasks. The patterns in missingness, discussed previously, lend strength to this hypothesis and potentially shed light on which students find these tasks difficult. However, the division of items that emerged was not expected, and is difficult to reconcile with our current interpretation of IBM. Overall, the current solution proposes a challenge to our theory, and although we will move forward with data collection using these items, more analyses will be required to tease apart their meanings.

Ultimately, the alignment items were dropped from the survey (indicated in bold in Table A5). This was for two reasons – first, one factor loaded with only two items, below the suggested threshold, and one item failed to load at all, suggesting a problematic underlying solution [23]. Second, these items repeated the pattern produced by the salient identity items, in which student tasks were grouped together and behaved unlike research tasks. Based on this information, and due to the need to consolidate the number of items and avoid redundancy, we decided to drop the alignment items.

Conclusion

In conclusion, these results suggest that the preliminary items will capture the constructs of interest within the graduate population, but some vigilance and additional investigation will still be required. As patterns in the missing data indicate that items were not skipped at random, follow-up interviews and surveys will be needed to explore what caused students to overlook or ignore these items. Before analyzing data gathered by the final survey, a missing data analysis should be done to confirm whether the pattern remains, and a sensitivity analysis may need to be done in order to evaluate how the missing data would impact research conclusions. Additionally, although a need for counterbalancing is not indicated, as emotional state was not impacted by survey content, it is worthwhile to continue pursuing this topic in planned follow-up interviews, to further reduce the possibility of measurement error.

Finally, while factor analysis was used to explore and refine the list of included items, the population is still a new one and so further exploratory work may be required. The final list of items was selected to maximize parsimony and focus on constructs relevant to graduate population, but future results may indicate that other approaches are better. The results reported here contribute to the existing literature by outlining some of the ways that engineering identity and identity-based motivation are experienced differently by graduate students, but without comparative testing between the two groups strong conclusions cannot be reached. Overall, these results indicate the potential for new hypotheses regarding the differences between graduate and undergraduate populations, and in-depth exploration of how identities shift and evolve as students transition from one space to the other.

References

[1] President's Council of Advisors on Science and Technology, "Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics," *Science (80-.).*, vol. 2, p. 130, 2012.

[2] B. E. . 1960- Lovitts, *Leaving the ivory tower : the causes and consequences of departure from doctoral study*. Lanham : Rowman & Littlefield, 2001.

[3] R. Sowell, J. Allum, and H. Okahana, *Doctoral initiative on minority attrition and completion*. Washington, DC, 2015.

[4] A. Godwin, "The Development of a Measure of Engineering Identity," *123rd Am. Soc. Eng. Educ. Annu. Conf. Expo.*, p. 15, 2016.

[5] J. C. Hilpert, J. Husman, G. S. Stump, W. Kim, W. T. Chung, and M. A. Duggan, "Examining students' future time perspective: Pathways to knowledge building," *Jpn. Psychol. Res.*, vol. 54, no. 3, pp. 229–240, Sep. 2012.

[6] S. Malcom and M. Feder, Eds., *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees*. Washington, D.C.: National Academies Press, 2016.

[7] Institute of Medicine, National Academy of Science, and National Academy of Engineering, *Reshaping the Graduate Education of Scientists and Engineers*. Washington, D.C.: National Academies Press, 1995.

[8] S. Jones, "More than an intervention: strategies for increasing diversity and inclusion in STEM," *J. Multicult. Educ.*, vol. 10, no. 2, pp. 234–246, 2016.

[9] A. I. Leshner, "Rethinking graduate education.," *Science (New York, N.Y.)*, vol. 349, no. 6246, American Association for the Advancement of Science, p. 349, 24-Jul-2015.

[10] C. Cass *et al.*, "Improving performance and retention of engineering graduate students through motivation and identity formation," in *ASEE Annual Conference and Exposition, Conference Proceedings*, 2017, vol. 2017–June.

[11] H. Perkins, M. A. Tsugawa-Nieves, J. N. Chestnut, B. Miller, A. Kirn, and C. Cass, "The role of engineering identity in engineering doctoral students' experiences," in *ASEE Annual Conference and Exposition, Conference Proceedings*, 2017, vol. 2017–June.

[12] M. A. Tsugawa-Nieves, H. Perkins, B. Miller, J. N. Chestnut, C. Cass, and A. Kirn, "The role of engineering doctoral students' future goals on perceived task usefulness," *ASEE Annu. Conf. Expo. Conf. Proc.*, vol. 2017–June, 2017.

[13] B. Miller, M. A. Tsugawa-Nieves, J. N. Chestnut, H. Perkins, C. Cass, and A. Kirn, "The influence of perceived identity fit on engineering doctoral student motivation and performance," in *ASEE Annual Conference and Exposition, Conference Proceedings*, 2017, vol. 2017–June.

[14] D. Oyserman, "Identity-Based Motivation," in *Emerging Trends in the Social and Behavioral Sciences*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2015, pp. 1–11.

[15] L. Benson *et al.*, "Characterizing Student Identities in Engineering: Attitudinal Profiles of Engineering Majors," *Am. Soc. Eng. Educ. Annu. Conf. Expo.*, Jun. 2017.

[16] B. Weijters, E. Cabooter, and N. Schillewaert, "The effect of rating scale format on response styles: The number of response categories and response category labels," *Int. J. Res. Mark.*, vol. 27, no. 3, pp. 236–247, Sep. 2010.

[17] K. C. Elmore and D. Oyserman, “If ‘we’ can succeed, ‘I’ can too: Identity-based motivation and gender in the classroom.,” *Contemp. Educ. Psychol.*, vol. 37, no. 3, pp. 176–185, Jul. 2012.

[18] E. R. Thompson, “Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS),” *J. Cross. Cult. Psychol.*, vol. 38, no. 2, p. 227-242--, 2007.

[19] D. Watson, L. A. Clark, and A. Tellegen, “Development and Validation of Brief Measures of Positive and Negative Affect - the Panas Scales,” *J. Pers. Soc. Psychol.*, vol. 54, no. 6, pp. 1063–1070, 1988.

[20] D. Wilson *et al.*, “Belonging and Academic Engagement Among Undergraduate STEM Students: A Multi-institutional Study,” *Res. High. Educ.*, vol. 56, no. 7, pp. 750–776, 2015.

[21] B. L. Yoder, “Engineering by the Numbers,” 2015.

[22] R. J. A. Little, “A Test of Missing Completely at Random for Multivariate Data with Missing Values,” *J. Am. Stat. Assoc.*, vol. 83, no. 404, p. 1198, Dec. 1988.

[23] A. B. Costello and J. W. Osborne, “Best Practices in Exploratory Factor Analysis : Four Recommendations for Getting the Most From Your Analysis,” *Pract. Assessment, Res. Educ.*, vol. 10, no. 7, pp. 1–9, 2005.

[24] I. Izquierdo, J. Olea, and F. Abad, “Exploratory factor analysis in validation studies: Uses and recommendations.,” *Psicothema*, vol. 26, no. 3, pp. 395–400, 2014.

[25] K. J. Preacher and R. C. MacCallum, “Repairing Tom Swift’s Electric Factor Analysis Machine,” *Underst. Stat.*, vol. 2, no. 1, pp. 13–43, 2003.

[26] L. G. Grimm and P. R. Yarnold, *Reading and understanding multivariate statistics*. American Psychological Association, 1995.

[27] G. Raîche, T. A. Walls, D. Magis, M. Riopel, and J.-G. Blais, “Non-Graphical Solutions for Cattell’s Scree Test,” *Methodology*, vol. 9, no. 1, pp. 23–29, Jan. 2013.

[28] E. Zell, Z. Krizan, and S. R. Teeter, “Evaluating gender similarities and differences using metasynthesis,” *Am. Psychol.*, vol. 70, no. 1, pp. 10–20, 2015.

[29] A. R. Taylor, B. D. Lutz, C. Hampton, W. C. Lee, and B. A. Watford, “Critical pedagogies and first-year engineering students’ conceptions of ‘what it means to be an engineer,’” *ASEE Annu. Conf. Expo. Conf. Proc.*, vol. 2017–June, 2017.

[30] F. L. Smyth and B. A. Nosek, “On the gender-science stereotypes held by Scientists: Explicit accord with gender-ratios, implicit accord with scientific identity,” *Front. Psychol.*, vol. 6, no. MAR, 2015.

[31] M. H. Lin, V. S. Y. Kwan, A. Cheung, and S. T. Fiske, “Stereotype content model explains prejudice for an envied outgroup: Scale of anti-Asian American Stereotypes.,” *Personal. Soc. Psychol. Bull.*, vol. 31, pp. 34–47, 2005.

[32] T. Brown, *Confirmatory factor analysis for applied research*, Second Edi. The Guilford Press, 2015.

Appendix

Table A1. Scientist Identity Survey Items, Uniquenesses, Communalities, and Factor Loadings

Item	Uniquenesses	Communalities	Recognition	Performance/ Competence	Interest
My peers see me as a SCIENTIST	0.300	0.781	0.883		
My advisor(s) see me as a SCIENTIST	0.292	0.837	0.849		
Other scientists see me as a SCIENTIST	0.318	0.835	0.812		
My department faculty see me as a SCIENTIST	0.449	0.667	0.778		
I have had experiences in which I was recognized as a SCIENTIST	0.532	0.617	0.734		
My family sees me as a SCIENTIST	0.573	0.581	0.712		
I see myself as a SCIENTIST	0.427	0.834	0.657		
Others ask me for help with SCIENCE	0.550	0.762	0.52		
I want to be recognized for my contributions to SCIENCE	0.642	0.712	0.354		
I am confident that I can understand SCIENCE in class	0.278	0.834		0.852	
I understand concepts I have studied in SCIENCE	0.313	0.768		0.849	
I can perform well when my SCIENCE knowledge is tested (for instance, in exams or defenses)	0.435	0.728		0.83	
I am confident that I can understand SCIENCE outside of class	0.272	0.874		0.809	
I can overcome setbacks when learning SCIENCE	0.492	0.751		0.614	
I am interested in learning more about SCIENCE in my field	0.094	0.958			0.923
I enjoy learning SCIENCE	0.109	0.963			0.873
I find satisfaction when learning more about SCIENCE	0.224	0.911			0.812

Note: Items indicated in bold were dropped from final survey.

Table A2. Engineering Identity Survey Items, Uniquenesses, Communalities, and Factor Loadings

Item	Uniquenesses	Communalities	Recognition	Performance/ Competence	Interest
Others ask me for help with ENGINEERING	0.246	0.86	0.904		
My peers see me as an ENGINEER	0.307	0.771	0.898		
I have had experiences in which I was recognized as an ENGINEER	0.392	0.776	0.812		
Other engineers see me as an ENGINEER	0.292	0.837	0.761		
My family sees me as an ENGINEER	0.555	0.661	0.636		
My advisor sees me as an ENGINEER	0.336	0.852	0.609		
I see myself as an ENGINEER	0.322	0.932	0.410		0.415
I understand concepts I have studied in ENGINEERING	0.170	0.866		0.936	
I am confident I can understand ENGINEERING outside of class	0.213	0.879		0.891	
I am confident I can apply ENGINEERING to solve problems	0.316	0.779		0.885	
I can perform well when my ENGINEERING knowledge is tested (for instance, in exams or defenses)	0.352	0.793		0.792	
I am confident I can understand ENGINEERING in class	0.350	0.823		0.784	
I can overcome setbacks when doing ENGINEERING	0.526	0.842		0.397	0.317
I enjoy learning ENGINEERING	0.298	0.868			0.932
I find satisfaction when doing ENGINEERING	0.466	0.785			0.798
I want to be recognized for my contributions to ENGINEERING	0.162	0.834			0.792
My department faculty see me as an ENGINEER *	0.280	0.737	1.010		
I am interested in learning more about ENGINEERING *	0.005	0.977			1.132

Note: Items indicated in bold were dropped from final survey.
* Item was a Heywood case (factor loading > 1) and dropped. Analyses were re-run, producing values for other items reported here.

Table A3. Researcher Identity Survey Items, Uniquenesses, Communalities, and Factor Loadings

Item	Uniquenesses	Communalities	Recognition	Performance/ Competence	Interest
My department faculty see me as a RESEARCHER	0.167	0.895	0.958		
My peers see me as a RESEARCHER	0.087	0.959	0.945		
Other researchers see me as a RESEARCHER	0.119	0.947	0.888		
My advisor(s) see me as a RESEARCHER	0.231	0.895	0.811		
I have had experiences in which I was recognized as a RESEARCHER	0.221	0.938	0.685		
My family sees me as a RESEARCHER	0.328	0.832	0.607		0.396
Others ask me for help with RESEARCH	0.271	0.922	0.592		
I see myself as a RESEARCHER	0.226	0.956	0.564		
I am confident that I can discuss current research trends	0.279	0.823	-0.319	0.839	0.303
I can publish research results in my field	0.155	0.899		0.957	
I am confident that I can design a RESEARCH study	0.273	0.839		0.873	
I can obtain my desired RESEARCH results	0.301	0.818		0.858	
I understand the concepts needed to analyze and interpret data	0.289	0.858		0.801	
I can present research related topics to relevant audiences	0.345	0.812		0.786	
I am confident that I can network with other researchers	0.346	0.813		0.75	
I am confident that I can find and understand literature in my field	0.387	0.8		0.738	
I can apply current industry trends and practices in my RESEARCH	0.582	0.629		0.683	
I understand how to use the necessary tools/equipment for my research	0.424	0.79		0.669	
I am confident that I can work with other researchers	0.306	0.888		0.636	
I am confident in my ability to mentor junior members of my research group	0.467	0.775		0.582	

I can overcome setbacks when doing RESEARCH	0.218	0.978		0.362	0.398
I enjoy conducting RESEARCH	0.154	0.909			0.954
I find satisfaction when doing RESEARCH	0.144	0.95			0.852
I find satisfaction when learning about my RESEARCH topic	0.224	0.911			0.839
I am interested in learning more about how to do RESEARCH	0.287	0.828			0.809
I want to be recognized for my contributions to RESEARCH	0.279	0.823			0.469

Items indicated in bold were dropped from final survey.

Table A4. Identity-Based Motivation Items, Uniquenesses, Communalities, and Factor Loadings (Salient Identity)

Item	Uniquenesses	Communalities	Scientist	Engineer	Researcher	Student
When I read journal articles, I feel like a(n) scientist	0.220	0.849	0.974			
Overall, I identify as a(n) scientist	0.201	0.753	0.962			
When I conduct research, I feel like a(n) scientist	0.195	0.955	0.926			
When I write peer-reviewed papers, I feel like a(n) scientist	0.327	0.854	0.805			
When I collaborate with other graduate students, I feel like a(n) scientist	0.388	0.771	0.770			
When I present my results, I feel like a(n) scientist	0.275	0.841	0.664			
When I attend conferences, I feel like a(n) scientist	0.376	0.864	0.542			
Overall, I identify as a(n) engineer	0.265	0.701		0.917		
When I collaborate with other graduate students, I feel like a(n) engineer	0.376	0.895		0.737		
When I read journal articles, I feel like a(n) engineer	0.543	0.591		0.722		
When I conduct research, I feel like a(n) engineer	0.487	0.735		0.662		
When I write peer-reviewed papers, I feel like a(n) engineer	0.609	0.537		0.633		
When I attend conferences, I feel like a(n) engineer	0.497	0.785		0.630		
When I present my results, I feel like a(n) engineer	0.459	0.814		0.630		
When I attend classes, I feel like a(n) engineer	0.646	0.678		0.475		0.325
When I present my results, I feel like a(n) researcher	0.184	0.975			0.883	
When I attend conferences, I feel like a(n) researcher	0.318	0.904			0.739	
When I read journal articles, I feel like a(n) researcher	0.297	0.876			0.730	
Overall, I identify as a(n) researcher	0.376	0.772			0.697	
When I write peer-reviewed papers, I feel like a(n) researcher	0.407	0.790			0.574	
When I conduct research, I feel like a(n) researcher	0.251	0.879			0.548	
When I collaborate with other graduate students, I feel like a(n) researcher	0.420	0.752	0.369		0.468	

When I do homework, I feel like a(n) researcher	0.199	0.750			0.969
When I do homework, I feel like a(n) scientist	0.261	0.772			0.846
When I attend classes, I feel like a(n) researcher	0.335	0.880			0.787
When I attend classes, I feel like a(n) scientist	0.336	0.919	0.389		0.638
When I do homework, I feel like a(n) engineer	0.730	0.671			0.339

Table A5. Identity-Based Motivation Items, Uniquenesses, Communalities, and Factor Loadings (Task Alignment and Difficulty)

Item	Uniquenesses	Communalities	Aligned Research Tasks	Aligned Student Tasks	Difficult Communication Tasks	Difficult Production Tasks
Alignment: Presenting my research	0.227	1.056	0.891			
Alignment: Conducting research	0.254	0.805	0.838			
Alignment: Writing about my research	0.374	0.683	0.779			
Alignment: Reading research publications	0.434	0.758	0.749			
Alignment: Attending conferences	0.487	0.856	0.724			
Alignment: Completing coursework/homework	0.053	0.930		0.981		
Alignment: Attending classes	0.14	0.834		0.921		
Difficulty: Writing about my research	0.134	1.063			0.950	
Difficulty: Presenting my research	0.399	0.977			0.724	
Difficulty: Reading research publications	0.466	0.568			0.715	
Difficulty: Attending classes	0.422	0.680				0.789
Difficulty: Completing coursework/homework	0.614	0.683				0.614
Difficulty: Collaborating with other graduate students (writing/research)	0.627	0.546				0.518
Difficulty: Conducting research	0.788	0.479				0.435
Difficulty: Attending conferences	0.809	0.420				0.331
Alignment: Collaborating with other graduate students (writing/research)	0.721	0.443				

'Alignment' items prefaced by text, "The following tasks align with how I view myself . . .".
'Difficulty' items prefaced by text, "I consider the following tasks to be difficult . . ."

Table A6. Regression Analysis of Items with Data MNAR (Missing Not At Random)

Item	Demographic	t	p	β	F	df	p	adj. R^2
When I write peer-reviewed papers, I feel like a engineer	International Status	1.773	0.079	0.235				
	Race/Ethnicity	0.264	0.792	0.035				
	Gender *	3.082	0.003	0.275				
	Overall Model				4.657	3, 116	0.004	0.084
When I write peer-reviewed papers, I feel like a scientist	International Status **	2.100	0.038	0.274				
	Race/Ethnicity	0.223	0.824	0.029				
	Gender *	3.381	> .001	0.296				
	Overall Model				6.061	3, 116	> .001	0.113
When I write peer-reviewed papers, I feel like a researcher	International Status	1.770	0.079	0.234				
	Race/Ethnicity	0.691	0.491	0.092				
	Gender *	3.375	0.001	0.300				
	Overall Model				4.959	3, 116	0.003	0.091
Overall, I identify as a scientist	International Status	0.009	0.993	0.001				
	Race/Ethnicity ***	-2.402	0.018	-0.322				
	Gender	0.616	0.538	0.055				
	Overall Model				4.335	3, 116	0.006	0.078

* Item skipped significantly more often by male students
** Item skipped significantly more often by international students
*** Item skipped significantly more often by Asian students