

Investigating Engineering Identity Development and Stability Amongst First-Year Engineering Students: A Person-Centered Approach

Mary Elizabeth Lockhart, Corresponding Author, Department of Teaching, Learning & Culture, Texas A&M University, 4232 TAMU College Station, TX, 77843, USA, mlockhart@tamu.edu

Karen Rambo-Hernandez, Department of Teaching, Learning & Culture, and Educational Psychology, Texas A&M University, 4232 TAMU College Station, TX, 77843, USA
rambohernandez@tamu.edu

Abstract

Engineering identity is a rapidly evolving construct within engineering, primarily due to its link to engineering persistence. Yet, most engineering identity research has been qualitative in nature or has described individuals within the analytical technique as coming from a single population. This study is the first to allow for the detection of different meaningful groups of engineering students demonstrating similarity on the construct using the new technique of Random Intercept Latent Transition Analysis. Through this study we identified three stable classes of engineering identity existed amongst first-year undergraduate students. Women demonstrated a greater likelihood of advancing to higher engineering identity classes over time than men. Unfortunately, the influence of COVID-19 yielded negative engineering identity developmental patterns for some students. Lastly, descriptive analyses of students' first-year engineering identity class assignments in relation to their selection/non-selection of engineering majors revealed Calculus-readiness upon college entrance might be an important component in these relationships.

Keywords

Engineering identity, engineering education, Random Intercept Latent Transition Analysis, person-centered

Introduction

The National Academies Gathering Storm committee concluded several years ago that the primary driver of the future economy, security of the United States (US) as a nation, and concomitant creation of jobs would be innovation—largely derived from advances in science and, particularly, in engineering (National Research Council, 2007). The European Union (EU) has mirrored this notion regarding the importance of engineering to the success and longevity of its related countries by prioritizing the production of a sufficient number of engineering graduates prepared to enter engineering professions over the last several decades (Gómez et al., 2022). Yet, both North America and Europe struggle with attrition in engineering. It has been found that almost half of students in the US and an estimated 30-40% of students in European countries such as Spain who begin their studies in engineering, do not reach graduation (Araque et al., 2009; García-Ros et al., 2019; Litzler & Young, 2012). These rates are alarming.

Engineering identity is a rapidly evolving construct within the field of engineering. It has been found to be a strong predictor of student retention and success in engineering degree programs and fields (Beam et al., 2009; Pierrakos et al., 2009; Rodriguez et al., 2018; Tonso, 2014). Thus, cross-disciplinary research teams from engineering, education, psychology, and sociology have been urgently seeking to define and investigate the construct (Revelo et al., 2019; Rodriguez et al., 2018). Yet to date, most of the research has been qualitative in nature (Rodriguez et al., 2018) or has described individuals within the analytical technique as coming from a single population. These studies, though noteworthy, have not allowed for the detection of different, meaningful groups, or classes, of individuals demonstrating similarities in their engineering identities (Jung & Wickrama, 2008).

Using person-centered analytical approaches can provide tremendous insight into engineering identity development, cultivation, stability over time, factors influencing its stability, and ultimately its relationship to students' persistence into the engineering workforce—as have been documented with other identity domains (Luyckx et al., 2008b; Meeus et al., 2012). This study is a first to apply such techniques to investigate the stability of engineering students' engineering identity classifications over their first year in college, the potential impact of gender and COVID-19 on this stability, and the relationship between these classifications and students' selection or non-selection of engineering majors at the beginning of their second year in college. Though based in the US, the culmination of this study serves to advance the international research base regarding engineering identity by providing unique insights into the construct's development and cultivation across different groups of engineering students that can only be obtained through person-centered analytical approaches.

Engineering Identity

Theoretical perspectives regarding identity formation were first initiated by Erikson. Erickson (1959, 1968) believed that identity was a primary task of adolescence that resulted from individuals beginning to cope with social and developmental demands while, simultaneously, seeking to provide meaning to their life choices and commitments (Bosma & Kunnen, 2008; Hewlett, 2013; Jensen, 2011; McLean & Syed, 2015; Schwartz et al., 2011; Was et al., 2009). Adolescents must make important decisions in multiple identity domains, such as education and interpersonal relationships, that ultimately lead to identity synthesis and cultivation or identity crisis (Albareello et al., 2017; Branje et al., 2014; McLean et al., 2016).

Following Erickson, many studies have further proposed that an individual can be classified under different identity statuses depending upon which identity domain is under

consideration (Marcia et al., 1993). Indeed, advances in identity theory, social identity theory, and research on identity in practice have come to recognize that people have multiple identity classifications as they simultaneously serve multiple roles in their lives (Stryker & Burke, 2000; Tajfel & Turner, 1986). A woman engineer, for example, might also be a researcher, daughter, mother and musician; thus, her identity might be developing differently depending upon the identity domain being measured.

One identity domain gaining the attention of researchers and practitioners alike resides in the field of engineering. Engineering identity has quickly risen to the frontlines of investigations in engineering, drawing upon cross-disciplinary research from engineering, education, psychology, and sociology (Revelo et al., 2019; Rodriguez et al., 2018). In actuality, the conceptualization of engineering identity can be traced back to Gee (2000), a linguist, who attempted to provide a bridge from traditional identity theory posited by Erickson (1959) into education. Gee loosely defined identity as a “kind of person” one is in any given context. Many recent studies seeking to measure engineering identity have been built upon the grounded model of science identity put forward by Carlone and Johnson (2007) that utilizes Gee’s definition of identity (Revelo et al., 2019). Carlone and Johnson suggested the existence of three interrelated dimensions of science identity: Competence, Performance, and Recognition. Inside of Carlone and Johnson’s description of the Recognition dimension of science identity resonates Gee’s definition of identity, or recognizing oneself as a “science kind of person” (Carlone & Johnson, 2007; Chemers et al., 2011; Gee, 2000; Godwin, 2016; Godwin et al., 2013).

Tracking this theoretical lens from science identity into the domain of engineering identity, several scholars have applied the grounded theory of science identity established by Carlone and Johnson (2007) to the field of engineering (Melo et al., 2017; Godwin, 2016;

Godwin et al., 2013; Revelo et al., 2019; Rodriguez et al., 2018). They have, thus, simultaneously carried forward Gee's (2000) definition of identity as being a "kind of person" into the field of engineering. It follows that one of the central components of a student's engineering identity is largely recognized as being the measurement of the degree to which they "view themselves as an engineering kind of person." Godwin (2016) briefly elaborated on this operationalization and defined engineering identity as a student's ability to feel like the kind of person who is interested in, possesses the relevant knowledge and skills in, and engages in engineering practices. University students with strong engineering identities establish and refine their engineering interests, build competence within the area of engineering, and utilize various tools to help them perform and cultivate their engineering identities (Godwin, 2016; Godwin et al., 2013).

As of late, engineering identity has been brought into the discussion regarding contributors to student retention and academic success in engineering, along with their ultimate matriculation into engineering fields (Bowen et al., 2019; Tonso, 2014; Rodriguez et al., 2018). Student GPA at the end of their first year in college, SAT math scores, ACT math scores, high school GPA and Calculus-readiness—being eligible to take a Calculus course their first semester in college—have all been identified as significant variables that impact student retention within engineering programs (Bowen et al., 2019; French et al., 2005; Hall et al., 2015; Lam et al., 1999; Levin & Wyckoff, 1988). Indeed, the importance of mathematical achievement and preparation to engineering persistence has long been studied. Correlations between first semester undergraduate mathematics course grades and success/persistence in engineering have been recorded, regardless of which mathematics course was taken (Budny et al., 1998). Moreover, the significant predictive nature of students' Calculus-readiness upon college entrance

to their actual attainment of an engineering degree have also been well documented with Calculus-ready engineering students being significantly more likely to graduate with a degree in engineering than non-Calculus-ready engineering students (Bowen et al., 2017; Bowen et al., 2019). However, more recent, the construct of engineering identity has joined this list of potential influences on engineering persistence.

Prior research has suggested that students who do not identify with the engineering field have a greater likelihood of exiting or not entering the engineering workforce than those who do (McCave et al., 2014; Owen & Rolfes, 2015; Rodriguez et al., 2018; Tonso, 2014; Trytten et al., 2015). Meyers and colleagues (2012) suggested that undergraduate engineering students must navigate acquiring necessary skills and disciplinary development for their careers as they begin cultivating their professional engineering identities. This cultivation of undergraduate students' engineering identities is imperative for positive experiences and successful retention within engineering degree programs leading into engineering fields (Meyers et al., 2012; Seymour, 1997; Stevens et al., 2005). Ultimately, successful cultivation of students' engineering identities impacts their continued pursuit of engineering majors and subsequent entrance into the engineering workforce (Beam et al., 2009; Pierrakos et al., 2009; Revelo et al., 2019). However, no knowledge currently exists regarding how engineering identity is related to engineering major selection or underperformance in engineering degree programs.

Given the amount of research suggesting the importance of engineering identity to the pursuit of engineering careers, and the importance of engineering professionals to the success and prosperity of the US as a nation (National Research Council, 2007), the establishment and cultivation of a strong engineering identity within undergraduate engineering students is critical. This establishment is perhaps even more important for women than men.

Gender Differences in Engineering Identity

It is no secret that gender disparities exist within STEM fields. Progress in reducing gender inequality, however, has occurred within some STEM disciplines. As documented by the National Science Foundation's (2021) findings, 41% of all bachelor's degrees in the physical sciences, 38% in the earth sciences, and 42% in mathematics and statistics are now being awarded to women. However, the gap in degree attainment between men and women remains substantially wider in engineering, where only 22% of all bachelor's degrees are awarded to women (National Science Foundation, 2021). Women, thus, continue to be underrepresented within engineering.

According to a study by Eccles and Barber (1999) as cited by Chemers and colleagues (2011), underrepresented groups in STEM (e.g., women in engineering) who identify strongly with academic role identities demonstrate greater persistence to degree completion than underrepresented students who identify more strongly with their social identities such as ethnic/racial identities or gender identity. Hamlet and colleagues (2020) noted specifically that engineering identity formation plays a crucial role in the persistence and retention of these underrepresented students in engineering fields. Engineering identity, thus, is gravely important to increasing gender diversity within the field.

Women have been found to begin their engineering coursework with lower confidence in several areas including: their engineering knowledge, their ability to succeed in engineering, and their perceptions of the various contributions that engineers make to society (Besterfield-Sacre et al., 2001). These lower levels of confidence persist for women throughout their engineering schooling although they do not perform at a lesser academic standard than men (Besterfield-Sacre et al., 2001; Seymour & Hewitt, 1997). Indeed, women often leave science, mathematics

or engineering majors having attained academic scores equivalent to, or higher than men who are retained within these majors (Seymour & Hewitt, 1997). Despite demonstrated engineering academic ability and skills, a lack of identification or belonging with engineering has repeatedly been suggested as a motivating factor for women who leave engineering (Godwin & Potvin, 2017; Seymour & Hewitt, 1997; Sheppard et al., 2015).

Women have been found to be more aware of social pressures to conform to the norms of the professional engineering culture than men, likely due to their already existing deviation from the gender norm (Hamlet et al., 2020). Engineering students who are unable to conform to culturally accepted norms and values, such as for women in the men-dominated engineering environment, leave their programs of study early (Dryburgh, 1999). There is a sense of belonging within engineering that seems critical for women as they learn to navigate cultural norms within the field. Furthermore, this crucial sense of belonging has been found to potentially provide an added protective component to women's engineering identities (Hamlet et al., 2020). Women may more deeply internalize their fit in engineering to help withstand their peers' disapproval due to their violation of the gender norm within engineering (Hamlet et al., 2020).

Attaining this deeper internalization of their engineering identity, however, is not easy and requires a great deal of effort by these women in the men-dominated engineering programs (Faulkner, 2007; Hamlet et al., 2020). A lack in this internalization of their engineering identity leads to women not persisting within the field (Godwin & Potvin, 2017; Seymour, 1997; Sheppard et al., 2015). Indeed, this internalization of engineering identity for women is critical to their persistence.

Current Research Practices

Though research regarding engineering identity development has increased substantially over the past 15 years, most of the study has been qualitative in nature (Rodriguez et al., 2018). Identity development examines how individuals progress through identity stages, or classifications, over time (Erickson, 1959, 1968; Marcia et al., 1993). This theoretical framework provides a rich soil for longitudinal, quantitative investigations.

Conventional longitudinal modeling approaches often assume that individuals come from a single population and that a single growth trajectory can adequately approximate that entire population (Jung & Wickrama, 2008). Person-centered quantitative methods, however, are especially applicable to modeling identity development (e.g., engineering identity development) as they are capable of detecting different, meaningful groups, or classes, of individuals demonstrating similarities on the construct (Jung & Wickrama, 2008; Luyckx et al., 2008b; Meeus et al., 2012). These classes can then be investigated over time for probabilistic transitions of individuals between classes and predictors of such transitions, which yields tremendous insight into identity development.

A few investigations have applied person-centered techniques to the study of a similar construct—science identity (Benedict et al., 2019; Lockhart, 2021; Robinson et al., 2018; Robinson et al., 2019; Shaby et al., 2020). Robinson and colleagues (2018, 2019) utilized a person-centered approach and found three distinct science identity classes of science college majors existed, two with flat development trajectories and the lowest class with a negative trajectory over time. These findings have led to insights regarding science identity development and best intervention practices for college students (Robinson et al., 2018; Robinson et al., 2019). No person-centered quantitative approach, however, has been utilized for engineering identity research. Thus, much insight is left to be gained regarding engineering identity

development over time within engineering students or factors contributing to, or hindering, its cultivation.

Present Study

Considering the evidence regarding the importance of engineering identity to engineering persistence for undergraduate students, this study examined the stability of engineering students' engineering identity over their first year in college. This study used the new longitudinal, person-centered approach of Random Intercept Latent Transition Analysis, RI-LTA (see *Data Analytic Strategy*), to classify first-year engineering students within engineering identity statuses and examine their transitions between statuses over an academic year. To date, no longitudinal, person-centered approach has been applied to engineering identity research in the published literature, nor has RI-LTA been utilized for any published study except the original work of Muthén and Asparouhov (2020). This study is, thus, a first for both.

In the current study, RI-LTA was utilized to investigate the existence of different engineering identity classes within first-year engineering majors and how likely students are to transition to different classes during the year. Gender and COVID-19 year were also investigated within this study for their potential influences on student probabilistic transitions to different engineering identity classifications throughout the academic year. Though the impact of COVID-19 upon college-aged individuals will not likely be fully recognized for many years, if ever, recent studies have documented that the instability, isolation, and strict online environments for college students during the first year-and-a-half of the pandemic adversely affected their physical and mental health (Chang et al., 2021; Castaneda-Babarro et al., 2020; Lopez-Moreno et al., 2020; Wang et al., 2020). It is important, therefore, that potential effects of

the pandemic upon students' engineering identity stability be investigated and relayed to the research community.

The following primary research questions (RQ) were addressed in this study:

- RQ1: Are there different engineering identity classifications for first-year undergraduate engineering students?
- RQ2: Are first-year undergraduate engineering students' engineering identity statuses, or classifications, stable over an academic year?
 - RQ2.1: Does gender influence engineering identity classification stability?
 - RQ2.2: Did COVID-19 influence engineering identity classification stability?
- RQ3: How are first-year engineering students' initial engineering identity classifications related to their selection or non-selection of engineering majors at the beginning of their second year?

Methods

Participants and Procedures

This study resided within a larger federal research grant study focused on changing undergraduate engineering curriculum and activities. Participants for this study were all first-year engineering students from a major, public, research university in the US that were enrolled in two consecutive introductory engineering classes over one academic year. All students were preliminarily accepted into the engineering program at the university and set to declare an engineering major at the beginning of their second year, as was common practice for this university. The study was reviewed and approved by the Institutional Review Board (IRB# 1905584259) and participant consent was obtained through the online survey platform.

Students' engineering identities were measured on three different occasions during the academic year (T1 = August, T2 = December, and T3 = May). These measurements were taken for three consecutive, distinct cohorts. Information was gathered and used in this study from a total of 598 out of 610 student participants who completed at least two of the three measurement occasions and were enrolled in both the fall and spring engineering courses. Five students were not included in this study due to missing data on the gender variable. The modeling technique utilized in this study would automatically exclude these five students from analyses utilizing gender as a covariate. Data from an additional seven students was also excluded from this study due to a gender selection of "other." Although important to capture, this group was too small to accurately represent within the analysis. Thus, data from 598 participants remained and is described below.

Cohort 1 (2018-2019) included 205 students with 31% women and 4% underrepresented minorities (URM). Cohort 2 (2019-2020) contained 294 student participants with 25% women and 7% URM. Cohort 3 (2020-2021) included 99 students with 33% women and 12% URM. Missing data was detected on at least one of the three measured variables (see *Measure*) across the three timepoints for a total of 38% of participants including approximately 1% of participants at T1, 13% of participants at T2, and 25% of participants at T3. The statistical technique used in this study afforded the opportunity to include all 598 participants in the analysis. Analyses were conducted to assess differences between those who participated at all three timepoints and those who did not. Mann-Whitney U Tests using the Bonferroni correction revealed that there were no significant differences between the two groups on any of the three measured identity variables (see *Measure*), at any of the three timepoints. These results suggested that data may be missing at random (Little & Rubin, 1987).

Measure

Identity as an Engineer. The Identity as a Scientist instrument developed by Chemers and colleagues (2010) was adopted and modified specifically for engineering. The extent to which participants identified as an engineer was measured using Chemers and colleagues' (2010) identity items. Through prior work with the adapted scale, the full instrument was repeatedly found to produce sub-adequate model fit results under confirmatory factor analysis. A refined scale with three of the items was selected that yielded a saturated model fit with high, statistically significant factor loadings greater than .75. Items were rated on a scale of 1 (strongly disagree) to 7 (strongly agree). Participants indicated their level of agreement with the statements such as, "I have come to think of myself as an engineer." Please see Appendix A for a full list of items. A higher scale score indicated a greater degree of self-identification as an engineer. Cronbach's alpha was used to assess the internal consistency of identity items for each of the three measurement occasions (T1-T3). Results revealed good internal consistency of the instrument with Cronbach's $\alpha = .88, .89$, and $.92$ for all cohorts combined at T1, T2, and T3, respectively (Kline, 1999). When separated by cohorts, Cronbach's alpha per each timepoint was greater than or equal to .85.

Data Analytic Strategy

To assess the research questions and sub-questions, RI-LTA was utilized to examine how probabilistic classes of students (i.e., engineering identity statuses) vary in systematic ways over time. According to Muthén and Asparouhov (2020), the person-centered regular Latent Transition Analysis (LTA) technique is unnecessarily restrictive. Regular LTA is a single-level modeling approach. RI-LTA, alternatively, reflects a multilevel modeling approach of separating the between and within-subject variation (Muthén & Asparouhov, 2020). By considering time as

the within-level and student as the between-level, the latent class transitions are represented on the within-level (Muthén & Asparouhov, 2020). The between-level, thus, captures much of the variability across students which yields more accurate classifications (Muthén & Asparouhov, 2020). Though LTA and RI-LTA are not normally recommended for small sample sizes ($N < 500$), Muthén (2021) illustrated RI-LTA typically fits the data better than LTA when $N \geq 500$ and there are at least three measurement occasions. Furthermore, RI-LTA leads to more accurate estimates of the transition probabilities, reduces the probability of subjects staying in the same class, and reduces the need for Mover-Stayer modeling when compared to LTA (Muthén, 2021; Muthén & Asparouhov, 2020).

To begin investigating RQ1, RQ2 and related sub-questions, three primary processes were employed: basic model identification, model invariance testing, and covariate inclusion (Muthén, 2021). Appropriate steps to identifying a baseline LTA model would typically begin with a series of latent class analyses (LCA) to identify the optimal number of subgroups of people, or classes, being represented at each timepoint (Nylund, 2007). This strategy, however, does not work for the RI-LTA modeling framework since the models differ by a random intercept factor—yielding LCA nonapplicable (Muthén, 2021). The optimal number of latent classes to help answer RQ1 were, thus, investigated through the estimation of various RI-LTA models. Model fit indicators such as loglikelihood (LL), Akaike information criterion (AIC), Bayesian information criterion (BIC), entropy (classification accuracy), and class size were used to aid in the selection of the most appropriate model. Higher loglikelihood values, lower AIC and BIC values, entropy values closer to 1.00 with a .70 cutoff (Clark, 2010; Fonseca & Cardoso, 2007; Ramaswamy et al., 1993) and reasonable class sizes containing at least 5% of the sample (Shanahan et al., 2013) were used as indicators of a better model fit (Muthén, 2021).

After a baseline model was selected, the appropriateness of the lag 1 assumption (where each measurement occasion directly influenced the next measurement occasion) was tested. A lag 2 model (where the first measurement occasion directly influenced the third measurement occasion) was estimated and compared to the lag 1 model using global model fit indices.

Once the baseline model with the appropriate lag was confirmed, the validity of the model based upon the assumptions of invariance over time and across groups was investigated. For examining the time invariance assumption, one indicator was freed and allowed to vary across all time points. This method was repeated for each indicator. Global model fit indices for each of these models were compared to the fully invariant model to determine if partial or full invariance held across time.

To examine measurement invariance across groups (gender and cohort), a flexible modeling approach utilizing covariates was utilized (Muthén, 2021). Specifically, a direct effects model where the grouping variable acts as a covariate and directly influences the latent class variables and latent class indicators (see Figure 1) was compared to a main effects model where the grouping variable acts as a covariate and influences the latent class variables and the random intercept (see Figure 2). The selection of the main effects model as the better fitting model would indicate measurement invariance held across the grouping variable as the random intercept captured most of the measurement non-invariance that was time-invariant (Muthén, 2021; Muthén & Asparouhov, 2020).

After invariance was established, RQ2 was addressed by examining transition probabilities obtained through the validated, baseline model. Odds ratios greater than one with a significant, non-symmetric 95% confidence interval (centered at 1.00) indicated that the odds of

transitioning from a particular class to another at that timepoint were significantly greater than the odds of remaining within the same class.

After establishing and validating the baseline model, covariates for gender (man or woman) and group, as related to students' cohort year (Cohort 1 – pre-COVID-19: 2018-2019, Cohort 2 – onset of COVID-19: 2019-2020, or Cohort 3 – mid-COVID-19: 2020-2021) were incorporated into the model one at a time as a main effects model to assess the degree to which they influenced the transitions of individuals from one classification status to another. RQ2.1 and RQ2.2 were investigated through their respective covariate effects on the transition probabilities and related significant or nonsignificant odds ratios. Significant odds ratios signified that the odds of transitioning from one class to another were significantly different for one-level of the covariate (e.g., women) than the other (e.g., men).

Lastly, after confirming and validating the optimal RI-LTA model and subsequent covariate effects, students' posterior probabilistic engineering identity classifications were analyzed. Specifically, RQ3 was examined through a descriptive analysis of students' probabilistic engineering identity classifications at T1 in relation to their selected major at the beginning of their second year.

STATA 17.1 (StataCorp, 2021) was used for all descriptive and correlational studies. RI-LTA models were estimated with Mplus Version 8.7 (Muthén & Muthén, 1998 –2021) using the maximum likelihood estimation method with robust standard errors (MLR), the default estimator for RI-LTA. Missing data was handled using the full information maximum likelihood (FIML) method, default to Mplus. All non-nested models used for invariance testing across groups were compared using BIC values with lower values indicating a better model fit. Nested models used for model identification and invariance testing over time were compared using BIC and

loglikelihood values with lower BIC and higher loglikelihood values indicating a better model fit. Where appropriate, formal Chi-Square Difference Tests were applied using the Satorra-Bentler Correction at the standard $\alpha = .05$ significance level (Muthén, 2021). If BIC and loglikelihood values were not in agreement as to which model was the better fitting model, BIC was selected for the model comparison (Muthén, 2021; Muthén & Asparouhov, 2020).

Results

The present study was based on an identity measure rated on a 7-point Likert scale. LTA and RI-LTA typically reduce a study's rating to a minimal number of categories, often binary, due to computational complexity (Collins & Lanza, 2010; Ryoo et al., 2018). Noting the potential loss of information this would entail, it was decided to proceed with the commonly accepted practice of analyzing the present study's data based on a continuous rating scale (Johnson & Creech, 1983; Norman, 2010; Sullivan & Artino, 2013). The present study ($n=598$) analyzed under the RI-LTA framework contains an adequate sample size to yield reliable results with $T \geq 3$, and $N \geq 500$ (Muthén, 2021). Descriptive statistics and a correlation matrix for the study variables are provided in Appendix B.

Model Identification

RI-LTA models for the 2, 3, and 4-class solutions were produced. Results are provided in Table 1. Class sizes are provided for each RI-LTA model at every timepoint, ordered from the lowest engineering identity class with the lowest indicator intercepts to the highest engineering identity class with the highest indicator intercepts. All models converged. Class intercept estimates and standard errors of the different class RI-LTA models are provided in Appendix C.

RQ1: Are there different engineering identity classifications for first-year undergraduate engineering students? The different RI-LTA models were compared. The 3-class model demonstrated substantially lower AIC (12255.63), lower BIC (12409.40) and higher LL (-6092.81, number of free parameters (fp) = 35) than the 2-class solution (AIC = 12526.23, BIC = 12627.28, LL = -6240.12, fp = 23). This signified the superiority of the 3-class solution to the 2-class solution. Entropy for the 3-class solution (.74) was less than that of the two-class solution (.82), but still above the .70 cutoff. The class sizes for both solutions were adequate. The 3-class model was deemed superior to the 2-class model.

The 4-class model had one class that captured only 1% to 2% of participants on each of the three measurement occasions. This did not meet our established criterion of an adequate class size containing at least 5% of participants. The 4-class model was not chosen to serve as the baseline model as it failed to meet our established criterion for model selection. The 3-class model was, thus, selected to serve as the baseline model. Subsequently, three engineering identity classifications of first-year engineering students were discovered—Medium, Medium-High, and High. We designated the lowest class as “Medium” because their scores on the identity items were moderate in nature (see Appendix C).

A **lag 2**, 3-class model was tested (BIC = 12426.38, LL = -6088.51, fp = 39) and compared to the 3-class baseline model. With a lower BIC and borderline insignificant adjusted chi-square difference test ($X^2(4) = 9.75, p=.045$), the baseline model was retained based upon our established criterion for model selection.

Model Invariance Testing

To examine the time invariance assumption of the baseline model, one indicator was freed at a time and allowed to vary across all time points. Model results are provided in Table 2. The loglikelihood for the fully invariant, baseline model was slightly higher than the partially invariant models. These differences in loglikelihood values, however, were only significant between the baseline model and the Item 3 Noninvariant model (*adjusted* $X^2(6) = 19.10, p = .004$). The BIC for the fully invariant model (BIC = 12409.40) was less than that of each of the partially invariant models, including the Item 3 Noninvariant model (BIC = 12420.53). Thus, the fully invariant baseline model was retained based upon our established criterion for model selection.

To examine invariance across groups (gender and cohort), a direct effects model was compared to a main effects model using BIC. The gender main effects model produced a lower BIC (12444.27) than the related direct effects model (12469.85). Similarly, the cohort main effects model also produced a lower BIC (12465.41) than the corresponding direct effects model (12536.41). Thus, the main effects model for both gender and cohort were superior to their direct effects counterparts. Measurement invariance was upheld across gender and cohort under the RI-LTA framework.

RQ2: Are first-year undergraduate engineering students' engineering identity statuses, or classifications, stable over an academic year? After measurement invariance was established across time and groups, RQ2 was addressed. The most common class patterns based on the estimated model were as follows: 3-3-3 where students remained in the highest engineering identity class across all three timepoints (36%); 2-3-3 where students started in the Medium-High class at T1 and then transitioned to the High class by T2 and remained there for T3 (11%); and 2-2-2 where students remained in the Medium-High class across all three timepoints (10%).

Next, transition probabilities between the timepoints were observed from the validated baseline model (see Table 3). Transition probability odds ratios are provided in Appendix D. Students tended to stay within the same class over time, though some movement was detected as was found in examination of the class patterns with 11% of students demonstrating the 2-3-3 pattern. Indeed, the greatest likelihood of transitioning to a different class was derived from this pattern and occurred between T1 and T2. Students in Class 2 (Medium-High) had a 37% chance of transitioning to Class 3 (High). However, the odds of transitioning from Class 2 to Class 3 during this time were .69 times that of just staying in Class 2. These odds were non-significant (0.37, 1.26). In general, the probabilistic movement of students between classes was not statistically significant, or significantly less than the likelihood of students remaining in their same class. The classes, thus, demonstrated stability over time.

Covariate Inclusion

RQ2.1: *Does gender influence engineering identity classification stability?* To address **RQ2.1**, a binary gender variable representative of man or woman (man = 0, woman = 1) was incorporated into the model as a covariate in a main effects model to examine its effects on the probabilistic transitions of students between classes. The effect of gender on transition probability odds ratios was observed over each time period. These ratios are provided in Table 4. Notably, between T2 and T3 the odds of transitioning from Class 2 (Medium-High) to Class 3 (High) were 2.42 times greater for women than for men. This result was statistically significant (1.07, 4.71). During this same time period, the odds of transitioning from Class 3 (High) to Class 2 (Medium-High) was .45 times as likely for women than men. This result was also statistically significant (0.21, 0.94). Hence, gender did demonstrate influence on engineering identity stability over the academic year.

RQ2.2: *Did COVID-19 influence engineering identity classification stability?* To address **RQ2.2**, two dummy-coded covariates representative of the COVID-year (Cohort 2: Cohort 2 = 1, otherwise = 0 and Cohort 3: Cohort 3 = 1, otherwise = 0) were incorporated into the model as main effects models to examine its effects on the probabilistic transitions of students between classes. Cohort 1 was selected to serve as the reference group because it was the only group not affected by COVID-19. The effect of cohort classification on the transition probability odds ratios were observed over each time period. These ratios representing the effects of Cohort 2 and Cohort 3 are provided in Table 5 and Table 6, respectively. Over the first semester, students in Cohort 2 were 3.18 times more likely to transition from Class 1 (Medium) to Class 3 (High) between the beginning and middle of the academic year compared to students in Cohort 1. During this same time period, students in Class 2 (Medium-High) were 2.64 times more likely to transition to Class 3 (High) if they were in Cohort 2 compared to Cohort 1. These results were statistically significant. It should be noted that both of these probabilistic transitions occurred during the first semester of the 2019 academic year—before the COVID-19 outbreak directly impacted the US but was already making headlines. In examination of the effects of Cohort 3 on transitions, it was found that students were 2.97 times more likely to transition down from Class 3 (High) to Class 1 (Medium) during their second semester if they were in Cohort 3 (mid COVID) compared to Cohort 1 (pre COVID). This result was statistically significant. COVID-19 as modeled by Cohort-year, thus, did demonstrate a significant influence on classification stability.

Descriptive Analysis

Lastly, to address *RQ3: How are first-year engineering students' initial engineering identity classifications related to their selection or non-selection of engineering majors at the*

beginning of their second year? students' probabilistic engineering identity classifications at T1 were observed in conjunction with their major selections at the beginning of Year 2. A total of 77 student participants' major selections at the beginning of their first and second years were missing. The reason(s) for this are unknown but could be contributed to a myriad of sources—the students could have dropped out of college, enrolled in a different university, had an “undeclared” major, or experienced a computer error in data gathering. Thus, we cannot adequately describe these students. Major selections for the remaining 521 student participants can be seen in Figure 3 (see Appendix E for the related table).

Students residing in Engineering Identity Class 3 (High) comprised the greatest proportion of each engineering major (see Figure 3). This is not surprising given that all the students in this study entered college planning to major in engineering, and that Engineering Identity Class 3 constituted the majority of all students at T1. Chemical Engineering and Industrial Engineering majors showed the greatest difference in representations between Engineering Identity Class 2 (Medium-High) and Class 3 (High) (see Figure 3). Indeed, 3.25 times more students who majored in Chemical Engineering (N=26) were in Engineering Identity Class 3 compared to Class 2 (N=8). Also, 2.38 times more students who majored in Industrial Engineering were in Engineering Identity Class 3 (High, N=31) compared to Class 2 (Medium-High, N=13). Furthermore, though a total of 20 students selected an Electrical Engineering Major for their second year, none of these students were in Engineering Identity Class 1 (Medium). This was the only major that did not have a Class 1 student.

Of special interest regarding retention within engineering education are the groups of students who “switched” to another major for Year 2, who were declared “Not on Track” for Year 2 and could not select an engineering major, or who did “Not Return” for Year 2.

Approximately 28, or 5%, of the student participants selected a major outside of engineering for Year 2—16 students who were initially classified in Engineering Identity Class 3 (High), 11 in Engineering Identity Class 2 (Medium-High) and one in Engineering Identity Class 1 (Medium).

Noting the importance of Calculus-readiness to the retention of engineering majors (e.g., Bowen et al., 2019; French et al., 2005) combined with the importance of engineering identity to the retention of engineering majors (Beam et al., 2009; Pierrakos et al., 2009; Rodriguez et al., 2018; Tonso, 2014) along with the lack of published literature investigating the links between Calculus-readiness and engineering identity, we decided to conduct an exploratory analysis and investigate students' Calculus-readiness upon college entrance in regards to their engineering identity and major selection. Calculus-readiness upon college entrance was determined by the university. A student was declared "Calculus-ready" if they were deemed eligible to enter any Calculus course and "not Calculus-ready" if they were not eligible to enter a Calculus course upon college entrance. Approximately 37% of students who entered college not Calculus-ready successfully progressed into an engineering major by Year 2. In contrast, about 60% of "switchers," entered college not Calculus-ready. Of these "switchers" who were not Calculus-ready, 63% of them were initially classified in Engineering Identity Class 3.

Furthermore, approximately 9% of student participants were declared "Not on Track" at the beginning of Year 2 and were, thus, unable to declare an engineering major—26 students initially classified in Engineering Identity Class 3 (High), 16 in Engineering Identity Class 2 (Medium-High), and 3 in Engineering Identity Class 1 (Medium). Again, utilizing Calculus-readiness to further describe these students, it was found that 80% of the students declared "Not on Track" entered college not Calculus-ready. Moreover, looking more specifically at students

by their class, 92% of Engineering Identity Class 3 (High) students who were declared “Not on Track” were not Calculus-ready upon college entrance.

Lastly, eight out of the 521 students beginning on an engineering track in Year 1 did not return for Year 2—seven initially classified in Engineering Identity Class 3 (High) and one in Engineering Identity Class 1 (Medium). Approximately 75% of students who did “Not Return” were not Calculus-ready upon college entrance, including five out of the seven students initially classified in Engineering Class 3.

Discussion

This study is the first in peer-reviewed published literature to apply person-centered quantitative techniques to the longitudinal study of engineering identity. It is also the first to apply the newly established RI-LTA modeling framework to a relevant investigation outside of Muthén and Asparouhov’s (2020) original work. Noting the heightened attention and importance of engineering identity, we sought to investigate the number of engineering classes present within first-year engineering students, the stability of these identities over an academic year, the potential impact that gender and COVID-19 had on engineering identity stability, and the relationship of engineering identity to the selection or non-selection of engineering majors.

First, the establishment of a solid 3-class RI-LTA model was critical to this study as it demonstrated different classes of engineering identity existed for first-year engineering students. Given that no person-centered technique has been applied to the domain of engineering identity, this is a novel finding. This finding further emphasizes the need to not assume that individuals come from a single population when studying different identity domains, but to allow for the

detection of different, meaningful groups, or classes, of individuals demonstrating similarities on the construct (Jung & Wickrama, 2008; Luyckx et al., 2008b; Meeus et al., 2012).

The discovered 3-class RI-LTA model for engineering identity also mirrors the findings of Robinson and colleagues (2018, 2019). Using person-centered longitudinal methods, Robinson and colleagues (2018, 2019) also discovered 3-classes of science identity existed for undergraduate science majors. This 3-class engineering identity solution suggests that students do not all begin their collegiate tenure with a high-level of established engineering identity. Indeed, medium, medium-high, and high-levels of engineering identity were all detected (see *Appendix C*). Thus, distinctiveness was detected between students' engineering identities even though all the students were pursuing engineering. Simply grouping all students together for analytical purposes would have forfeited this valuable information and future investigations into relationships between these classes and external variables. Discovering these classes allows for further investigations into motivational, personality, and various psychosocial factors contributing to or hindering the effective cultivation of students' engineering identities over time.

The three engineering identity classes demonstrated stability throughout the academic year suggesting first-year engineering students typically remained within their same class. Again, these results were similar to Robinson and colleagues' (2018, 2019) science identity findings where two of three classes yielded no significant trajectory changes over time. The modest transitions in engineering identity that were detected suggested a greater likelihood of students transitioning in an upward trajectory, that is, transitioning to a higher engineering identity class instead of a lower one. These findings amplify the importance of the establishment of a strong engineering identity before college entrance.

Interestingly, women in this study were more likely to transition to higher engineering identity classifications and less likely to transition to lower classifications compared to men. This was evidenced by women being approximately 2.5 times more likely than men to transition from the Medium-High to High class during the latter part of the academic year. This significant result provides some statistical evidence regarding the internalization of women's engineering identities that previous studies have suggested (Faulkner, 2007; Godwin & Potvin, 2017; Hamlet et al., 2020; Seymour, 1997; Sheppard et al., 2015). Women who were first-year engineering students with a medium-high level of engineering identity appeared to grow in their identity and have some protective barrier against identity regression, or confusion, that was distinctively different from their men counterparts. This finding serves to enhance the research-base regarding women in engineering by providing statistical support to the importance of the internalization of women's engineering identity over time that differs from men.

Investigations into the myriad of effects the COVID-19 pandemic has had upon different populations within society will likely continue to be unraveled for the foreseeable future. The drastic rise in mental health struggles for university students during the pandemic is of particular importance. Studies documented increases in anxiety rates among undergraduate students by about 50% and depression rates by over 60% during the first year-and-a-half of the pandemic (Coakley et al., 2021; Czeisler, 2020; Fruehwirth et al., 2021; Lopez-Moreno et al., 2020). Social distancing measures, online courses, pandemic fatigue and several other factors could have potentially contributed to this trend. The growing influx of such mental health struggles for university students combined with first-year engineering students being towards the beginning of the formative period for their engineering identity is cause for concern and investigation. It is reasonable to question if greater numbers of first-year engineering students enrolled during the

pandemic transitioned to lower engineering identity statuses as the academic year and pandemic progressed in contrast to similar students who were enrolled prior to the pandemic. This study showed that indeed COVID-19 did influence engineering identity stability in a negative way. First-year engineering students were 2.97 times more likely to transition down from the highest engineering identity class to the lowest during their second semester if they were in Cohort 3 (mid-COVID) compared to Cohort 1 (pre-COVID). As time wore on for students enrolled during the 2020 academic year, their engineering identity showed significant negative developmental patterns compared to students before the pandemic. This only adds to the vast number of negative effects COVID-19 has had upon our world. Certainly, these results are noteworthy and warrant further investigation into the many effects of the COVID-19 pandemic including its impact upon students' various developing identities.

Given that persistence data commonly associated with engineering identity (i.e., graduation with an engineering degree or entrance into the engineering workforce) would not be available for several more years for all students in this study, it was decided to investigate students' initial engineering identity classification with their selection/non-selection of an engineering major at the beginning of Year 2. Through the descriptive analysis of initial identity classification in conjunction with the selection of an engineering major in Year 2, it was discovered that a fairly equivalent distribution of the identity classes across major selections existed. This is not surprising as students demonstrating greater levels of cultivation of their engineering identity would be expected to pursue engineering majors at greater proportions than others as a representation of their engineering persistence—ultimately culminating in their entrance into the engineering workforce (Cave et al., 2014; Owen & Rolfes, 2015; Rodriguez et al., 2018; Tonso, 2014; Trytten et al., 2015).

The finding that zero students from Class 1 (Medium) selected an Electrical Engineering major, and this being the only major without a Class 1 student is noteworthy. Electrical engineering is well-known for being one of, if not the most, math-intensive engineering majors as it relies heavily upon the fields of mathematics and physics—which is also mathematically-intensive. Combining this result with the findings that 63% of “switchers,” 92% of students declared “not on track” for Year 2, and 71% of students that did “not return” for Year 2 were classified in Engineering Identity Class 3 (High) but were not Calculus-ready upon college entrance suggests that there could be a link between mathematics (i.e., mathematics-readiness, mathematics appreciation, mathematics motivation, etc.) and engineering identity. It is feasible that some students enter college strongly identifying themselves as engineers but are unaware of the mathematical aptitude and strenuousness required for engineering degrees. However, more study is needed. Certainly, the link between students being Calculus-ready upon college entrance and their retention within engineering degree programs established in previous studies (e.g., Bowen et al., 2019; French, 2005) was also visible in this investigation with 57% of “switchers” and 80% of those declared “not on track” by Year 2 being not Calculus-ready.

Implications for Future Research

This study sets the groundwork for future investigations into engineering identity development for entering engineering students. Person-centered longitudinal, quantitative approaches are needed to continue investigating the development and stability of engineering identity throughout the college-tenure. One possible theoretical framework for which to approach future person-centered methodological studies regarding engineering identity development is through the lens of variation theory. Variation theory reflects a theory of learning and experience that explains various ways a learner might come to see, understand, or

experience a given phenomenon in a particular way, and why certain students in similar situations (e.g., classrooms, programs of study) might perceive concepts or constructs differently (Bussey et al., 2013; Orgill, 2012). Given that the current study quantitatively identified three engineering classes, variation theory would seek to understand “why” students who are all pursuing engineering are quantitatively grouped into these three distinct classes—reflecting variations in their perceived identification with engineering. Future investigations could consider utilizing variation theory in longitudinal person-centered quantitative investigations regarding engineering identity to further identify factors related to engineering identity cultivation.

In this study, stability in engineering identity was observed over students’ first year. However, this does not imply that stability continues. More study is needed to determine if stability is maintained throughout the college-tenure and how this impacts persistence into engineering fields. This will enable scholars and researchers to detect periods of potential change in engineering identity and allow for more directed and effective intervention approaches. It will also provide insight into how the different engineering identity classifications are related to engineering persistence as given by entrance into the engineering workforce.

Furthermore, the finding that most students who did not declare an engineering major for Year 2 were not-Calculus ready is of significance as it applied to students of all engineering identity classes. Noting that Calculus-readiness has already been linked to student retention within engineering programs and that engineering identity is currently gaining traction in investigating its importance to student matriculation into engineering fields, it is feasible that all these constructs are intertwined. A basis exists for future studies to investigate potential links

between Calculus-readiness and engineering identity, and how the relationship between these two constructs ultimately influences engineering persistence.

Limitations

There are limitations to the present study that require attention. The sample size for this study, though adequate, was not optimal and from only one university. A larger study with a more diverse sample obtained from varying institutions across the US is needed to validate the results of this study. Findings from this study lack generalizability.

Though the entropy value of the 3-class baseline model, .74, was above the .70 cutoff, values closer to 1.00 would demonstrate greater classification accuracy. This was likely due to a combination of the sample size and the complexity of the modeling technique. A larger sample size would likely produce higher entropy values and more consistent decisions when comparing global fit indices. Furthermore, the descriptive analysis portion of this study utilized the posterior probabilistic engineering identity classes at T1 produced by the validated baseline model. Formal testing, such as regression analysis, that utilize latent classifications are typically not recommended if the mixture model produced an entropy value less than .80 (Clark, 2010). This study was a preliminary investigation into engineering identity classifications, indeed the first of its kind. The goal of the descriptive analysis was simply to gain better insight into how the engineering identity classifications spread-out across the engineering majors and non-majors. No formal testing was utilized. Thus, we argued the information gained from the descriptive analysis should be presented to provide baseline information that can be utilized as a springboard for future studies. No firm conclusions should be drawn from this portion of the study.

Discrepancies in global model comparison and test results were also detected between the baseline model and the Item 3 Noninvariant model, the gender main effects model, and the cohort main effects model. The baseline model was continuously selected for model building purposes due to its lower BIC values. However, the discrepancies with its lower loglikelihood values provide reason for further consideration of the retention of the other models in future investigations with larger samples.

Furthermore, though this study provided some insight into the underrepresented group of women engineers and their related engineering identity, other **underrepresented** groups were not considered as they accounted for a small percentage of the participants. Future studies with larger samples may be able to explore patterns **by** ethnicity and other student background characteristics such as first-generation status. Not accounting for students' ethnicity potentially introduced some bias into this study and findings related to engineering identity stability.

Lastly, the measure of engineering identity used in this study reflects common operationalizations put forth by various scholars (Melo et al., 2017; Godwin, 2016; Godwin et al., 2013; Revelo et al., 2019; Rodriguez et al., 2018). However, we acknowledge that it lacks in its ability to measure the three dimensions of their engineering (science) identity proposed by Carlone and Johnson (2008)—Competence, Performance and Recognition. The instrument primarily measures the Self-Recognition component of the Recognition dimension of engineering identity. Though cited by Carlone and Johnson (2008) as being extremely critical to one's science identity, the measure of one's engineering self-recognition does not likely encompass the entirety of their engineering identity.

Conclusion

As engineering identity is a rapidly developing construct within the field of engineering, it is imperative that various analytical approaches be utilized to validate, confirm, or contradict theories and qualitative findings related to the construct. The findings of this study underscore that we can obtain three primarily stable engineering identity classes amongst first-year engineering students. The stability of these classes was statistically significantly affected by gender, providing support to previous research citing women might more deeply internalize their engineering identities than men. Unfortunately, engineering identity stability was also affected by COVID-19 adding to the growing body of literature regarding the negative impacts of COVID-19 upon our world. Moreover, the classes of engineering identity in conjunction with students' Calculus-readiness upon college entrance appear to be related to their non-selection of engineering majors in their second year of college. These findings should serve to stimulate and refine future investigations into the construct of engineering identity with a goal of uncovering the factors related to its successful cultivation within students over time and, ultimately, its influence on persistence into engineering careers. (8,724 words)

Acknowledgements

Any opinions, findings, conclusions, or recommendations herein are those of the authors and do not necessarily reflect those of NSF.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the U.S. National Science Foundation (NSF) [grant number 1725880].

References

- Albarello, F., Crocetti, E., & Rubini, M. (2017). I and us: A longitudinal study on the Interplay of personal and social identity in adolescence. *Journal of Youth Adolescence*, 47, 689-702. doi: 10.1007/s10964-017-0791-4
- Araque, F., Roldán, C., & Salguero, A. (2009). Factors influencing university drop out rates. *Computers & Education*, 53(3), 563-574.
- Beam, T., & Pierrakos, O., & Constantz, J., & Johri, A., & Anderson, R. (2009, June), *Preliminary Findings on Freshmen Engineering Students' Professional Identity: Implications for Recruitment and Retention* [Paper presentation]. Annual Conference & Exposition, Austin, Texas. 10.18260/1-2—5112
- Benedict, B., Verdin, D., Rohde, J. A., Brown, H., Baker, R., Thielmeyer, A., & Godwin, A. (2019). An Early Adaptation of Identity Trajectory to Understand the Identities of Undergraduate Engineering Students. *2019 IEEE Frontiers in Education Conference (FIE), Frontiers in Education Conference (FIE), 2019 IEEE*, 1–5.
<https://doi.org/10.1109/FIE43999.2019.9028365>
- Besterfield-Sacre, M., Moreno, M., Shuman, L., & Atman, C. (2001). Gender and ethnicity differences in freshmen engineering student attitudes: A cross-institutional study. *Journal of Engineering Education*, 477-489.
- Bosma, H. A., & Kunnen, E. S. (2008). Identity-in-context is not yet identity development-in-context. *Journal of Adolescence*, 31, 281–289.
<https://doi.org/10.1016/j.adolescence.2008.03.001>

- Bowen, B.D., Wilkins, J. L. M., & Ernst, J. V. (2019). How calculus eligibility and at-risk status relate to graduation rate in engineering degree programs. *Journal of STEM Education: Innovations and Research*, 19(5), 26-31.
- Branje, S., Laninga-Wijnen, L., Yu, R., & Meeus, W. (2014). Associations among school and friendship identity in adolescence and romantic relationships and work in emerging adulthood. *Emerging Adulthood*, 2, 6 –16. <https://doi.org/10.1177/2167696813515851>
- Budny, D., LeBold, W., Bjedov, G. (1998). Assessment of the Impact of Freshmen Engineering Courses. *Journal of Engineering Education*, 87(4), 405-411.
- Bussey, T. J., Orgill, M., & Crippen, K. J. (2013). Variation theory: A theory of learning and a useful theoretical framework for chemical education research. *Chemistry Education Research and Practice*, 14, 9-22.
- Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44(8), 1187–1218.
- Castaneda-Babarro, A., & Arbillaga-Etxarri, A., Gutierrez-Santamaría, B., Coca, A. (2020). Physical activity changes during COVID-19 confinement. *International Journal of Environmental Research and Public Health*, 17(18).
- Chang, J., Ji, Y, Li, Y., Pan, H., & Su, P. (2021). Prevalence of anxiety symptoms and depressive symptoms among college students during COVID-19 pandemic: A meta-analysis. *Journal of Affective Disorders*, 292, 242-254.

- Chemers, M. M., Syed, M., Goza, B. K., Zurbriggen, E. L., Bearman, S., Crosby, F. J., ... & Morgan, E. M. (2010). *The role of self-efficacy and identity in mediating the effects of science support programs* (No. 5). Technical Report.
- Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Sociological Issues*, 67, 469–491.
- Clark, S. L. (2010). *Mixture modeling with behavioral data*. [Unpublished doctoral dissertation] University of California, Los Angeles.
- Coakley, K. E., Le, H., Silva, S. R., & Wilks, A. (2021). Anxiety is associated with appetitive traits in university students during the COVID-19 pandemic. *Nutrition Journal*, 20(45).
- Collins, L., and Lanza, S. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York, NY: John Wiley & Sons.
- Czeisler, M., Lane, R. I., Petrosky, E., Wiley, J. F., Christensen, A., Njai, R., et al. (2020). Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States. *MMWR and Morbidity and Mortality Weekly Report*, 69(32), 1049–1057. <https://doi.org/10.15585/mmwr.mm6932a1>.
- Dryburgh, H. (1999). Work hard, play hard: Women and professionalization in engineering—Adapting to the culture. *Gender and Society*, 13(5), 664–682.
- Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research*, 14, 10–43.

- Erikson, E. H. (1959). *Identity and the life cycle*. New York, NY: Norton.
- Erikson, E. H. (1968). *Identity, youth, and crisis*. New York: Norton.
- Faulkner, W. (2007). Nuts and bolts and people' gender-troubled engineering identities. *Social Studies of Science*, 37(3), 331–356.
- Fonseca, J. R., & Cardoso, M. G. (2007). Mixture-model cluster analysis using information theoretical criteria. *Intelligent Data Analysis*, 11, 155-173. doi:10.3233/ida-2007-11204
- French, B. F., Immekus, J. C., & Oakes, W. C. (2005). An examination of indicators of engineering students' success and persistence. *Journal of Engineering Education*, 94(4), 419-425. <http://doi.org/10.1002/j.2168-9830.2005.tb00869.x>
- Fruehwirth, J. C., Biswas, S., & Perreira, K. M. (2021). The Covid-19 pandemic and mental health of first-year college students: Examining the effect of Covid-19 stressors using longitudinal data. *PLoS One*, 16(3), e0247999.
- García-Ros, R., Pérez-González, F., Cavas-Martínez, F., & Tomás, J. M. (2019). Effects of pre-college variables and first-year engineering students' experiences on academic achievement and retention: A structural model. *International Journal of Technology and Design Education*, 29, 915-928.
- Gee, J. P. (2000). Identity as an analytic lens for research in education. *Review of Research in Education*, 25, 99–125.
- Godwin, A. (2016). The development of a measure of engineering identity. [Paper presentation]. ASEE 123rd Conference & Exposition, New Orleans, June 26-29.

- Godwin, A., & Potvin, G. (2017). Pushing and pulling Sara: A case study of the contrasting influences of high school and university experiences on engineering agency, identity, and participation. *Journal of Research in Science Teaching*, 54(4), 439-462.
- Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2013). Understanding engineering identity through structural equation modeling. [Paper presentation]. IEEE Frontiers in Education Conference, Oklahoma City.
- Gómez, J., Tayebi, A., & Delgado, C. (2022). Factors that influence career choice in engineering students in Spain: A gender perspective. *IEEE Transactions on Education*, 65(1), 81-92.
Doi: 10.1109/TE.2021.3093655
- Hall, C. W., Kauffmann, P. J., Wuensch, K. L., Swart, W. E., DeUrquidí, K. A., Griffin, O. H. & Duncan, C. S. (2015). Aptitude and personality traits in retention of engineering students. *Journal of Engineering Education*, 104, 167-188. <https://doi.org/10.1002/jee.20072>
- Hamlet, L. C., Roy, A., Scalone, G., Lee, R., Poleacovschi, C., & Kaminsky, J. (2020). Gender and engineering identity among upper-division undergraduate students. *Journal of Management in Engineering*, 37(2).
- Hazari, Z., Wulff, P., Petersen, S., Neumann, K. (2018). Engaging young women in physics: An intervention to support young women's physics identity development. *Physical Review Physics Education Research*, 14(2).
- Hewlett, B. L. (2013). *Adolescent identity: Evolutionary, cultural and developmental perspectives*. New York, NY: Routledge.

- Jensen, L. A. (2011). *Bridging cultural and developmental approaches to psychology: New syntheses in theory, research, and policy*. New York, NY: Oxford University Press.
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302-317.
- Kline, P. (1999). *A handbook of psychological testing, 2nd edition*. London: Routledge.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Lam, P., Doverspike, D., & Mawasha, R. (1999). Predicting success in a minority engineering program. *Journal of Engineering Education*, 88, 265-2267. doi: 10.1002/j.2168-9830.1999. Tb00445
- Levin, J., & Wyckoff, J. (1988). Effective advising: Identifying students most likely to persist and succeed in engineering. *Engineering Education*, Vol. 78, 178-182.
- Litzler, E., & Young, J. (2012). Understanding the risk of attrition in undergraduate engineering: Results from the project to assess climate in engineering. *Journal of Engineering Education*, 101, 319-345.
- Lockhart, M. E. (2021). *A Potential Fix for the Leaky STEM Pipeline: Development and Validation of the SciID Scale*. (Unpublished doctoral dissertation). Texas A&M University. <https://oaktrust.library.tamu.edu/handle/1969.1/193142>
- Lopez-Moreno, M., Lopez, M. T. I., Miguel, M., & Garces-Rimon, M. (2020). Physical and psychological effects related to food habits and lifestyle changes derived from COVID-19 home confinement in the Spanish population. *Nutrients*, 12(11), 3445.

- Luyckx, K., Klimstra, T. A., Duriez, B., Van Petegem, S., & Beyers, W. (2013). Personal identity processes from adolescence through the late 20's: Age trends, functionality, and depressive symptoms. *Social Development, 22*, 701–721.
- Luyckx, K., Schwartz, S., Berzonsky, M., Soenens, B., Vansteenkiste, M., Smits, I., et al. (2008a). Capturing ruminative exploration: Extending the four-dimensional model of identity formation in late adolescence. *Journal of Research on Personality, 42*, 58–82.
- Luyckx, K., Schwartz, S. J., Goossens, L., Soenens, B., & Beyers, W. (2008b). Developmental typologies of identity formation and adjustment in female emerging adults: A latent class growth analysis approach. *Journal of Research on Adolescence, 18*, 595–619.
- Marcia, J. E., Waterman, A. S., Matteson, D. R., Archer, S. L., & Orlofsky, J. I. (1993). *Ego identity: A handbook for psychosocial research*. New York: Springer-Verlag.
- McCave, E.J., Gilmore, J., & Burg, K. (2014). *Engineering and science student preparedness for research: Exploring the connections between student identity and readiness for research* [Paper presentation]. American Society for Engineering Education Annual Conference, Indianapolis, IN.
- McLean, K. C., & Syed, M. (Eds.). (2015). *The Oxford handbook of identity development*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199936564.001.0001>
- McLean, K. C., Syed, M., & Shucard, H. (2016). Bringing identity content to the fore: Links to identity development processes. *Emerging Adulthood, 4*, 356–364.
<https://doi.org/10.1177/2167696815626820>

- Meeus, W., Van de Schoot, R., Keijsers, L., Schwartz, S. J., & Branje, S. (2012). Identity statuses as developmental trajectories: A five-wave longitudinal study in early-to-middle and middle-to-late adolescence. *Journal of Youth and Adolescence*, 41, 1008–1021.
- Melo, J. M., Benedict, B., Clements, R., Perkins, H., & Godwin, A. (2020). See me as an engineer: Understanding the role of language and multiple role identities on engineering students' identity trajectory. *IEEE Frontiers in Education Conference (FIE)*, 1-9, doi: 10.1109/FIE44824.2020.9274017.
- Meyers, K. L., Ohland, M. W., Pawley, A. L., Silliman, S. E., & Smith, K. A. (2012). Factors relating to engineering identity. *Global Journal of Engineering Education*, 14(1), 119–131.
- Muthén, B. (2021 February 28). *Using Mplus to do latent transition analysis and random intercept latent transition analysis* [Video file]. Retrieved from <https://www.youtube.com/c/MplusVideos>
- Muthén, B., & Asparouhov, T. (2020, November 23). Latent Transition Analysis with Random Intercepts (RI-LTA). *Psychological Methods*. Advance online publication. <http://dx.doi.org/10.1037/met0000370>
- Muthén, L. K., & Muthén, B. O. (1998-2021). *Mplus user's guide (eighth edition)*. Muthén & Muthén.
- National Research Council. (2007). *Rising Above the Gathering Storm*. Washington, DC: The National Academies Press.

- National Science Foundation. (2021). *Women, minorities, and persons with disabilities in science and engineering*. National Science Foundation. Retrieved from <https://ncses.nsf.gov/pubs/nsf21321/report/field-of-degree-minorities>
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), pp. 625-632. Retrieved from: <https://link.springer.com/article/10.1007%2Fs10459-010-9222-y#citeas>
- Nylund, K. L. (2007). *Latent transition analysis: Modeling extensions and an application to peer victimization* (Doctoral dissertation, University of California, Los Angeles). Retrieved from <https://www.statmodel.com/download/nylunddis.pdf>
- Orgill, M. (2012), Variation theory. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3391–3393). Heidelberg, Germany: Springer-Verlag GmbH.
- Owen, C. & Rolfes, D. (2015). *Communication class size and professional identity* [Paper presentation]. American Society for Engineering Education Annual Conference, Seattle, WA.
- Pierrakos, O., Beam, T. K., Constantz, J., Johri, A., & Anderson, R. (2009, October). *On the development of a professional identity: Engineering persisters vs engineering switchers* [Paper presentation]. Frontiers in Education Conference, San Antonio, TX.
- Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, 12, 103-124. doi:10.1287/mksc.12.1.103

- Revelo, R. A., Omitoyin, J., Cardona, M., Nazempour, R., & Darabi, H. (2019). *Engineering identity profiles of low-SES, high achieving incoming engineering students* [Paper presentation]. IEEE Frontiers in Education Conference (FIE), Cincinnati, OH.
- Robinson, K. A., Perez, T., Carmel, J. H., & Linnenbrink-Garcia, L. (2019). Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit. *Contemporary Educational Psychology*, 56, 180–192. <https://doi.org/10.1016/j.cedpsych.2019.01.004>
- Robinson, K. A., Perez, T., Nuttall, A. K., Roseth, C. J., & Linnenbrink-Garcia, L. (2018). From science student to scientist: Predictors and outcomes of heterogeneous science identity trajectories in college. *Developmental Psychology*, 54(10), 1977–1992. <https://doi.org/10.1037/dev0000567>
- Rodriguez, S. L., Lu, C., & Bartlett, M. (2018). Engineering identity development: A review of the higher education literature. *International Journal of Education in Mathematics, Science and Technology*, 6(3), 254–265.
- Ryoo, J. H., Wang, C., Swearer, S. M., Hull, M., Shi, D. (2018). Longitudinal model building using latent transition analysis: An example using school bullying data. *Frontiers in Psychology*, 9. doi: 10.3389/fpsyg.2018.00675
- Schwartz, S. J., Luyckx, K., & Vignoles, V. L. (2011). *Handbook of identity theory and research*. New York: Springer. <https://doi.org/10.1177/0894318411409427>
- Schwartz, S. J., Zamboanga, B. L., Wang, W., & Olthuis, J. (2009). Measuring identity from an Eriksonian perspective: Two sides of the same coin. *Journal of Personality Assessment*, 31, 143–154.

- Seymour, E., & Hewitt, N. M. (1997). *Talking about Leaving: Why Undergraduates Leave the Sciences*. Westview Press: Boulder, CO.
- Shaby, N., & Vedder, W. D. (2020). Science identity trajectories throughout school visits to a science museum. *Journal of Research in Science Teaching*, 57(5), 733–764.
<https://doi.org/10.1002/tea.21608>
- Shanahan, L., Copeland, W. E., Worthman, C. M., Erkanli, A., Angold, A., & Costello, E. J. (2013). Sex-differentiated changes in C-reactive protein from ages 9 to 21: The contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology*, 38(10), 2209-2217. <https://doi.org/10.1016/j.psyneuen.2013.04.010>
- Sheppard, S., Antonio, A., Brunhaver, S. & Gilmartin, S. (2015). Studying the career pathways of engineers: An illustration with two data sets. In *Cambridge Handbook of Engineering Education Research* (pp. 283–310). Cambridge, UK: Cambridge University Press.
- StataCorp. (2021). *Stata statistical software: Release 17*. StataCorp LLC.
- Stevens, R., O'Connor, K., & Garrison, L. (2005). Engineering student identities in the navigation of the undergraduate curriculum. In *Proceedings of the 2005 American Society for Engineering Education Annual Conference and Exposition*, 1–8. Reston, VA: ASCE.
- Stryker, S. & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social Psychological Quarterly*, 63, 284-297.
- Sullivan, G. & Artino Jr., A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*. 5(4), pp. 541-542.

- Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T. Jost & J. Sidanius (Eds.), *Political psychology: Key readings* (pp. 276–293). Psychology Press.
- Tonso, K. L. (2007). *On the outskirts of engineering: Learning identity, gender, and power via engineering practice*. Leiden, Netherlands: Brill Sense.
- Tonso, K. L. (2014). Engineering identity. In A. Johri & B. Olds (Eds.), *Cambridge handbook of engineering education research* (pp. 267-282). New York, NY: Cambridge University Press.
- Trytten, D. A., Pan, R., Shehab, R. L., & Walden, S. E. (2015). *Inclusion or exclusion? The impact of the intersection of team culture and student identity and pathway on team diversity* [Paper presentation]. American Society for Engineering Education Annual Conference, Seattle, WA.
- Wang, Y., Zhang, Y., Bennell, K., White, D. K., Wei, J., Wu, Z., He, H., Liu, S., Luo, X., Hu, S., Zeng, C., & Lei, G. (2020). Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: longitudinal observational study. *Journal of Medical Internet Research*, 22.
- Was, C. A, Al-Harthy, I., Stack-Oden, M., & Isaacson, R. M. (2009). Academic identity status and the relationship to achievement goal orientation. *Electronic Journal of Research in Educational Psychology*, 7(2), 627–652.

Appendix A

Engineering Identity Items

ID1: In general, being an engineer is an important part of my self-image.

ID2: Being an engineer is an important reflection of who I am.

ID3: I have come to think of myself as “an engineer”.

Note. Items adopted from: Chemers, M. M., Syed, M., Goza, B. K., Zurbriggen, E. L., Bearman, S., Crosby, F. J., ... & Morgan, E. M. (2010). *The role of self-efficacy and identity in mediating the effects of science support programs* (No. 5). Technical Report.

Appendix B

Descriptive Statistics and Correlation Matrix of Study Variables

	M	SD	ID1_ T1	ID2_ T1	ID3_ T1	ID1_ T2	ID2_ T2	ID3_ T2	ID1_ T3	ID2_ T3	ID3_ T3	Woman	Cohort 2	Cohort 3
ID1_T1	5.38	1.32	1.00											
ID2_T1	5.37	1.29	0.79	1.00										
ID3_T1	5.27	1.39	0.63	0.69	1.00									
ID1_T2	5.40	1.35	0.57	0.57	0.48	1.00								
ID2_T2	5.40	1.32	0.53	0.62	0.5	0.83	1.00							
ID3_T2	5.41	1.37	0.44	0.52	0.52	0.65	0.69	1.00						
ID1_T3	5.32	1.44	0.54	0.55	0.52	0.64	0.66	0.57	1.00					
ID2_T3	5.33	1.40	0.48	0.55	0.5	0.59	0.63	0.53	0.90	1.00				
ID3_T3	5.50	1.41	0.41	0.48	0.55	0.50	0.53	0.58	0.71	0.73	1.00			
Woman	0.29	0.45	0.02	-0.01	-0.02	-0.06	-0.03	-0.04	0.02	0.04	0.01	1.00		
Cohort 2	0.49	0.50	-0.04	-0.03	0.02	0.12	0.08	0.01	0.09	0.13	0.11	-0.05	1.00	
Cohort 3	0.17	0.37	-0.02	-0.03	-0.05	0.03	0.03	0.06	-0.02	-0.06	-0.06	.04	-0.39	1.00

Note. M = Mean, SD = Standard Deviation, T1 = Time 1, T2 = Time 2, T3 = Time 3, ID1 = *In general, being an engineer is an important part of my self-image*, ID2 = *Being an engineer is an important reflection of who I am*, ID3 = *I have come to think of myself as “an engineer”*.

Appendix C

Class Intercept Estimates and Standard Errors for the RI-LTA Models

	2 Class Solution		3 Class Solution			4 Class Solution			
Class Intercept Estimate (S.E.)/ Item Number	Class 1	Class 2	Class 1 (Medium)	Class 2 (Medium-High)	Class 3 (High)	Class 1	Class 2	Class 3	Class 4
ID1	3.58 (.17)	5.74 (.05)	2.70 (.20)	4.78 (.13)	6.08 (.09)	2.99 (.13)	2.68 (.30)	4.94 (.10)	6.02 (.06)
ID2	3.69 (.14)	5.72 (.05)	3.09 (.15)	4.74 (.11)	6.04 (.06)	2.96 (.11)	5.48 (.26)	4.77 (.09)	6.04 (.06)
ID3	4.23 (.17)	5.61 (.05)	3.97 (.16)	4.93 (.11)	5.83 (.06)	3.93 (.17)	5.04 (.33)	4.93 (.09)	5.84 (.07)

Note. Mean (standard deviation).

Appendix D

Transition Probability Odds Ratios for the 3-Class Model

		Time 2		
		Class 1	Class 2	Class 3
Time 1	Class 1	1.00 (1.00, 1.00)	0.66 (.23, 1.88)	0.59 (.24, 1.46)
	Class 2	0.19 (.09, .39)	1.00 (1.00, 1.00)	0.69 (.37, 1.26)
	Class 3	0.05 (.02, .105)	0.24 (.14, .41)	1.00 (1.00, 1.00)
		Time 3		
Time 2	Class 1	1.00 (1.00, 1.00)	0.62 (.24, 1.64)	0.26 (.07, .95)
	Class 2	0.24 (.11, .51)	1.00 (1.00, 1.00)	0.63 (.23, 1.77)
	Class3	0.03 (.01, .10)	0.19 (.09, .38)	1.00 (1.00, 1.00)

Note. Odds ratio (non-symmetric 95% confidence intervals centered at 1.00).

Appendix E

Engineering Identity Classifications by Engineering Major Table

Year2Major	IDT1			Total
	Class 1	Class 2	Class 3	
Aerospace Engineering	9 (9%)	38 (38%)	52 (52%)	99
Biomedical Engineering	3 (4%)	32 (46%)	35 (50%)	70
Biometric Systems Engineering	1 (33%)	0 (0%)	2 (67%)	3
Chemical Engineering	4 (11%)	8 (21%)	26 (68%)	38
Civil Engineering	6 (12%)	16 (32%)	28 (56%)	50
Computer Engineering	2 (6%)	11 (31%)	23 (64%)	36
Electrical Engineering	0 (0%)	8 (40%)	12 (60%)	20
Industrial Engineering	1 (2%)	13 (29%)	31 (69%)	45
Mechanical Engineering	7 (12%)	21 (36%)	31 (53%)	59
Mining Engineering	1 (10%)	1 (10%)	8 (80%)	10
Petroleum & Natural Gas Engineering	1 (10%)	1 (10%)	8 (80%)	10
Not Return	0 (0%)	1 (13%)	7 (88%)	8
Not on Track	3 (7%)	16 (36%)	26 (58%)	45
Switch	1 (4%)	11 (39%)	16 (57%)	28
Total	39	177	305	521

Note. IDT1 = Engineering identity measured at Time 1. Count (percent of major).