2017

Testing for Measurement Invariance in Engineering Identity Constructs for First-Generation College Students

Dina Verdín
Purdue University, dverdin@purdue.edu

Allison Godwin
Purdue University, godwina@purdue.edu

Testing for Measurement Invariance in Engineering Identity Constructs for First-Generation College Students

Dina Verdín, Allison Godwin
School of Engineering Education
Purdue University
West Lafayette, Indiana 47907
dverdin@purdue.edu, godwina@purdue.edu

***Abstract*—This work-in-progress research study examines the response patterns of first-generation college students (FGCS) to the engineering identity measures compared to non-first-generation college students (non-FGCS). This work answers the following research question, "Do FGC and non-FGC engineering students interpret the engineering identity measurement items in a conceptually different manner?" We explore if FGCS respond to engineering identity items similarly to non-FGCS and the fairness of using these instruments for FGCS to make claims about this group. The data for this work are from a survey instrument completed by 2,916 first-year engineering college students from four U.S. institutions. We hypothesize that quantitative measures constructed for the general engineering student population (non-FGCS) may not function the same for a FGCS subpopulation in engineering. Using extensions to the confirmatory factor analysis, we tested for measurement invariance of engineering identity constructs between FGCS and non-FGCS. Our comparative analysis of FGCS and non-FGCS found weak measurement invariance within the engineering identity constructs (i.e. interest, recognition, and performance/competence) indicating a similar factor structure and factor loadings, but different uses of the identity item scale. This research raises questions on the use and fairness of normative measures in engineering education for populations that fall outside the majority engineering student population.**

Keywords—first-generation college students; engineering identity; confirmatory factor analysis

I. Introduction

The process of how students take on an engineering identity has been used to understand important outcomes like learning in engineering context as well as students' pathways and persistence in engineering [1]. The act of identifying or becoming an engineer is important to students' navigation of their engineering curricula and integration within students' respective engineering programs [1],[2]. Studies have found that having a strong engineering identity contributes to retention and persistence [3], [4], specifically as it relates to underrepresented students [5], [6]. A recent study by Godwin and colleagues [7] examined students' mathematics and physics identities with three constructs: interest in the subject, recognition by others, and beliefs about one's ability to do well in a course (performance) and understand the course material (competence). The study used a large-scale, national dataset to conduct factor analysis and structural equation modeling to examine the validity of the measures used as well as the relationships between the factors in predicting engineering career choice. In a later study, the same author provided strong validity evidence for similar measures of engineering identity [8]. These studies explored the measurement of engineering identity, in general, but did not examine if students from diverse backgrounds interpreted the questions the same as their majority counterparts. This present study is a first-step in examining how a diverse group of students, first-generation college students, not only differ in terms of lived experiences [9], but also in the way they identify as a an engineer.

First-generation college students (FGCS) navigate the system of higher education without insider knowledge of the system from their parents, family members, or even siblings. That is, FGCS come from a family where neither parent has attended or completed a post-secondary education, while non-first-generation college students (non-FGCS) report having at least one parent complete post-secondary education [10]. We chose to focus on FGCS in this study because they are not likely to have direct experience with a family member who is an engineer, unlike their non-FGCS peers [11]. Because of this lack of connection with other engineers prior to college, these students may interpret items measuring how they see themselves as the kind of people who do engineering differently. Literature on engineering identity has focused broadly on the larger student body, with few emphases placed on a subpopulation of the engineering student body i.e., FGCS. It is important to understand how all students identify as engineers and more specifically, how diverse group of students' identification as an engineer may be different than their peers. We examine whether FGCS respond to engineering identity items similarly to non-FGCS and the implications on the fairness of using these instruments for FGCS to make claims about this group.

II. Theoretical Framework

How students identify with a particular STEM subject has been conceptualized through a subject-related role identity framework developed by Hazari and colleagues [12] who sought to describe students' identification as a physics person. Within

This material is based upon work supported by the National Science Foundation under Grant No. (1428689 and 1428523). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

their framework were characteristics relating to how students identify with a STEM discipline, these characteristics include interest in the subject, recognition (i.e., by peers, parents, and teachers), and performance/competence (i.e., student's ability to achieve good grades and their ability to understand concepts) [7], [8]. The interrelationship between the constructs have been articulated in relation to science identity in that "a satisfactory science identity hinges not only upon having competence and interest in science, but also critically, upon recognition by others as someone with talent and potential in science" [13, p. 1197]. We apply this understanding of science identity, as it relates with the three constructs, to student's identification with engineering. Prior literature has supported the use of these constructs to understand how students begin to identify as a science person [14], physics person [15] mathematics person [5], and as an engineer [8]. In this paper, we focus on understanding FGCS responses to engineering identity measures of the constructs of interest, recognition and performance/competence.

III. Research Question

By examining the response patterns of first-generation college students (FGCS) on the engineering identity measures compared to non-first-generation college students (non-FGCS). We sought to answer the following research question:

Do FGC and non-FGC engineering students interpret the engineering identity constructs in a conceptually different manner?

We hypothesize that quantitative measures constructed for the general engineering student population may not function the same for a FGCS subpopulation in engineering.

IV. Method

The data for this study came from a survey administered in Fall 2015 at three land-grant institutions and one Hispanic-Serving Institution. The population for this study was first-year engineering students yielding a sample of $N = 2{,}916$. The survey was administered via paper-pencil format and was completed during the first two weeks of classes. Providing paper-pencil format of the survey ensured high response rates [16]; these surveys were later digitized for analysis. The survey comprised of a set of items measuring students' attitudinal profiles including measures of belongingness in engineering, STEM identities (i.e., engineering, physics, and math), other affective measures, and demographic information (i.e., parent(s) level of education, race/ethnicity, gender) as well as students' career goals and choice of engineering major. In this study, we focus on the items measuring students' engineering identity for measurement invariance testing.

A. Measures

Students' response to a question about their parent/guardian level of education for either parent/guardian with "bachelor's degree" or "master's degree or higher," were coded as 0 = non-FGCS, whereas students' responses indicating both parent/guardian level of education "less than a high school diploma," "high school diploma/GED," or "some college or associate/trade degree," were coded as 1 = FGCS. Students who did not report parent's education level were eliminated from the study as we could not determine their status. Our definition of first-generation students is consistent with the U.S. Department of Education's classification [17].

The other items used in this analysis, engineering identity, were measured on a seven-point anchored numeric scale (0 = "Strongly Disagree" to 6 = "Strongly Agree"). Each construct comprised of three subject-related measures (i.e., performance/competence, interest, and recognition). The items measuring engineering identity have been previously published [8].

V. Analytic Approach

Measurement invariance is concerned with identifying whether a construct (e.g., engineering identity) has the same factor structure and interpretation for different groups. Engineering education researchers often report the results of factor analysis to demonstrate construct validity; however, factor analysis automatically assumes outcome variables are equivalent across groups unless explicitly tested. We tested the underlying assumptions for factor analysis at four increasingly restrictive levels to determine if FGCS and non-FGCS respond equivalently on measures of engineering identity. A multiple group confirmatory factor analysis (CFA) is a widely used method for explicitly testing measurement equivalence/invariance and entails a simultaneous analysis of a measurement model for FGCS and non-FGCS [18]. The absence of measurement invariance indicates that the results of pairwise comparisons of FGCS and non-FGCS on these measures cannot be isolated from differences in the group responses to the items. Therefore, the items cannot be used to infer differences among students as written [15]. In this work-in-progress study we offer a brief examination of the following type of measurement invariance: (Model 1) configural invariance, (Model 2) metric invariance, (Model 3) scalar invariance, and (Model 4) strict invariance. These four levels of invariance are tested in a stepwise fashion with increasingly rigorous assumptions to determine where differences, if any, occur for group response patterns to the engineering identity items. Each model and its assumptions are explained below. Data were analyzed using the R programming language and statistical software [20], and tests for measurement invariance was conducted through the lavaan package [21].

A. Model 1: Configural Invariance

Measurement invariance testing starts with configural invariance. Configural invariance tests whether there is an equal factor structure, that is, "the number of factors and pattern of indicator-factor loadings are identical across groups" [18, p. 242-3]. Configural invariance involves specifying a Confirmatory Factor Analysis (CFA) that reflects how the construct is theoretically operationalized. In our case, this process involves specifying which measurement items load onto each underlying engineering identity construct. This CFA-model is fitted separately for each group (i.e., FGCS and non-FGCS) and is examined to see if the theoretical structure is valid in each group. To continue testing for measurement invariance, configural invariance must be established [18], [19]. Once it is established that the basic structure of the model holds for both groups, testing for metric invariance (equivalence of factor loadings across groups) can be conducted.

B. Model 2: Metric Invariance

A test for quality of factor loadings is referred to as metric invariance. Metric invariance is performed after configural invariance is supported between the two groups. Metric invariance tests whether respondents across groups attribute the same meaning to the latent constructs under study (i.e., interest, recognition, and performance/competence beliefs). Testing for metric invariance uses a chi-square difference test to establish if constraining the factor loadings to be equal across the multiple groups corresponds with a significant increase in chi-square. A significant increase in chi-square results in a significant decrease in model fit, while a non-significant chi-square would support metric invariance [19].

C. Model 3: Scalar Invariance

Scalar invariance examines equality of intercepts between engineering identity scores across groups. Scalar invariance suggests that different groups respond to the scale (seven-point anchored numeric scale) in the same way. That is, a student identified as an FGCS and another student identified as a non-FGCS with the same level on the factor should obtain the same score on the seven-point anchored numeric scale [19]. Rejecting scalar invariance would suggest that (1) group differences in estimated factor means are biased and (2) group differences from the mean (generated from the measurement scale) or the estimated factor scores will not be directly related to the factor means and will be distorted by differential additive response styles [22]. A differential additive response style "occurs when one group systematically gives higher or lower responses than another group, resulting in a scale displacement" [23, p. 190]. Differential additive bias will inevitably make the mean differences of the observed variables smaller or larger than their true mean difference or will indicate no difference when a difference exists [18]. If scalar invariance is established, groups can be compared on their scores on the latent variable.

D. Model 4: Strict Invariance

Strict invariance assumes there is scalar invariance and tests for equality in error variances and covariance across FGCS and non-FGCS. This test examines if the residual (uniqueness or measurement error) associated with each measurement variable, the factor loadings of the latent variables, and the intercepts of the measured variables is equal across FGCS and non-FGCS [22]. Strict invariance indicates that the variance not explained by the model are different between groups and can result in unfair mean comparisons. When strict invariance is found, the differences between FGCS and non-FGCS item responses are solely due to group differences. The absence of strict invariance indicates an apparent item bias [22].

VI. Results

Of the students who participated in the survey, 72% (n_1 = 2,092) were classified as non-FGCS, 20% (n_2 = 596) FGCS, and 8% (n_3 = 228) did not report parental education status. Students who did not report parental education status were removed from this analysis. The engineering identity items (i.e., interest, recognition, and performance/competence) used in this study had high internal consistency with Cronbach's α for all measured variables above 0.80.

A. Model 1: Configural Invariance

Configural invariance was examined by testing the original three-factor structure for engineering identity [8]. In this test, no equality constraints were place, that is, all parameters were freely estimated for FGCS and non-FGCS separately. To assess the adequacy of fit for both FGCS and non-FGCS models, χ^2 tests and goodness-of-fit indexes were used. However, χ^2 is sensitive to large sample sizes therefore three fit indexes were used as additional evidence for model fit. Table I describes the three-factor solution fit for FGCS (CFI = .970, RMSEA = .033) and non-FGCS (CFI = .981, RMSEA = .071), providing verification of configural invariance across both student groups. The fit indexes indicated that the factor structure fit well for both groups or configural invariance. This result provides justification for conducting a multiple group CFA to test for model invariance. The three-factor solution for first-generation college students and non-first-generation college students is depicted in Fig. 1.

TABLE I. Fit Indexes for the Three-Factor Model of Engineering Identity Scale Across Two Groups

Group	χ^2	*df*	CFI	RMSEA	SRMR	*n*
FGCS	123.544***	36	0.970	0.086	0.033	587
non-FGCS	264.302***	23	0.981	0.071	0.029	2,062

***p<.001

TABLE II. Summary of Fit Statistics for Testing Measurement Invariance of Three-Factor Model of Engineering Identity

Fit Index	χ^2	*df*	Model Comparison	$\Delta\chi^2$	Δdf	ΔP	CFI	RMSEA	SRMR
Model 1: Configural Invariance	379.14	46					.98	.08	.04
Model 2: Metric Invariance (loadings)	388.37	52	2 vs. 1	9.23	6	0.16	.98	.07	.04
Model 3: Scale Invariance (intercepts)	402.13	58	3 vs. 2	13.76	6	0.03*	.98	.07	.03
Model 4: Strict Invariance (means)	423.73	61	4 vs. 3	21.60	3	0.00***	.98	.07	.03

Notes: RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual, CFI = comparative fit index, $\Delta\chi^2$ = likelihood ration test (chi-square difference test)

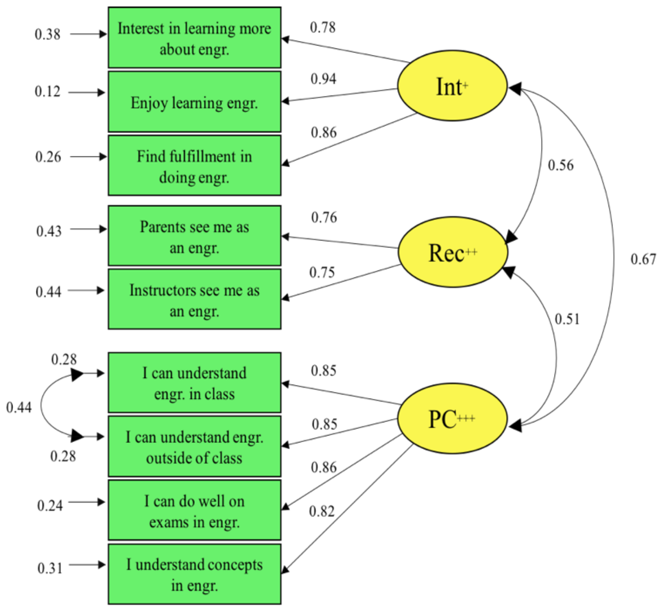

Fig. 1. Confirmatory factor analysis of engineering identity latent constructs
Notes: +Interest, ++Recognition, +++Performance/Competence

B. Model 2: Metric Invariance

Metric Invariance assumes configural invariance. Testing for metric invariance constrained the factor loadings to be equal across FGCS and non-FGCS response scales. The model fit here was found to be acceptable ($\chi^2(52) = 388.37$, $p < .001$; CFI = .98; RMSEA = .07; SRMR = .04) and did not significantly differ from the baseline configural model ($\Delta\chi^2(6) = 9.23$, $p = .16$) as shown in Table II indicating that the restriction of making the groups equal across groups did not significantly affect the fit of the model for the data. From this model, we concluded that the factor loadings could be estimated simultaneously for each group and that weak invariance assumptions (both configural and metric invariance) were met. Weak invariance indicates that the model has the same structure and same underlying constructs of interest, recognition, and performance/competence beliefs for FGCS and non-FGCS.

C. Model 3: Scalar Invariance

The next model tested Scalar Invariance or that the intercepts of the models estimated independently for the two groups was no better than the model estimated simultaneously for the two groups. The requirements for scalar or strict invariance were not met as indicated by chi-square difference tests between model with equality constrained factor loadings and intercepts and the metric invariance model as shown in Table II. In our tests, the chi-square difference tests between Model 2 and Model 3 are significantly different with $p = .03$ indicating that Model 3 with the constrained intercepts fit the data significantly worse than Model 2 and that intercepts should be estimated differently among FGCS and non-FGCS. This result indicates that the two groups respond to the given scales for the engineering identity items differently and that any comparisons of the composite scores on the constructs will be biased.

VII. Discussion

Our comparative analysis of FGCS and non-FGCS found weak measurement invariance for the engineering identity items. That is, while the overall structure and factor loadings of the subject-related identity constructs were consistent across groups, the analysis demonstrated a rejection of strong invariance indicating differences in students' use of the measurement scale. This variance in intercepts points to differences in the way FGCS are interpreting and answering identity measures in comparison to non-FGCS. These results indicate that using the construct scores for the engineering identity items to compare FGCS and non-FGCS would not be a fair use of this instrument. Any pairwise test or other comparisons of factor means will reflect differences not on true mean scores difference but on how students responded differently to the given scales. That is, students may have a different conceptual understanding of the Likert-scale. One example may be FGCS are more inclined to have indicate neutral responses (3 = Neither agree nor disagree), while non-FGCS may tend to respond towards the extreme ends (0 = "Strongly Disagree" or 6 = "Strongly Agree") or vice versa. Our future work will explore how FGCS and non-FGCS use the seven-point anchored numeric scales to provide recommendations to the engineering education community on how these items can be used to understand student differences.

VIII. Conclusion

Our work indicates that weak, but not strong (or strict) invariance exists for FGCS and non-FGCS on the engineering identity items used in this work. This work illustrates the need for group measurement invariance testing in addition to construct validity before items can be fairly used to compare groups. This work-in-progress raises questions for how engineering identity can be measured across diverse groups and understood to improve the quality of engineering education for all.

IX. Limitations & Future Work

This work-in-progress paper is a first step to understanding students' from different groups response patterns to measures of engineering identity scale. Missing from this small-scale study is an analysis of measurement invariance for different demographics (e.g., gender identity, race/ethnicity, institution) in both FGCS and non-FGCS groups. Similarly, an analysis of measurement invariance for students enrolled in different engineering disciplines may be conducted. Future work will seek to investigate measurement invariance between gender and engineering disciplines.

Acknowledgment

The authors acknowledge the InIce research team for the contributions in data collection and cleaning including Dr. Geoff Potvin, Dr. Lisa Benson, Dr. Adam Kirn, Dr. Monique Ross, Jacqueline Doyle, Hank Boone, Jacqueline Rohde, and Beverly Ma. We also thank the anonymous participants of the survey for the time taken to respond.

REFERENCES

[1] D. Verdin and A. Godwin, “Exploring Latina First-Generation College Students’ Multiple Identities, Self-efficacy, and Institutional Integration to Inform Achievement in Engineering,” *J. Women Minor. Sci. Eng.*, 2017.

[2] K. L. Tonso, “Teams that Work: Campus Culture, Engineering identity, and Social Interactions,” *J. Eng. Educ.*, vol. 95, no. 1, pp. 25–37, 2006.

[3] T. Beam, O. Pierrakos, J. Constantz, A. Johri, and R. Anderson, “Preliminary Findings on Freshmen Engineering Students ’ Professional Identity : Implications for Recruitment and Reten,” in *American Society for Engineering Education*, 2009.

[4] O. Pierrakos, T. K. Beam, J. Constantz, a Johri, and R. Anderson, “On the development of a professional identity: engineering persistors vs. engineering switchers.,” *Front. Educ. Conf.*, vol. 39, no. 2, pp. 599–604, 2009.

[5] C. E. Foor, S. E. Walden, and D. A. Trytten, “‘ I Wish that I Belonged More in this Whole Engineering Group :,’” *J. Eng. Educ.*, vol. 96, no. 2, pp. 103–115, 2007.

[6] R. A. R. Alonso, “Engineering Identity Development of Latina and Latino Members of the Society of Hispanic Professional Engineers Engineering Identity Development of Latina and Latino Members of,” 2015.

[7] A. Godwin, G. Potvin, Z. Hazari, and R. Lock, “Identity, Critical Agency, and Engineering: An Affective Model for Predicting Engineering as a Career Choice,” *J. Eng. Educ.*, vol. 105, no. 2, 2016.

[8] A. Godwin, “The Development of a Measure of Engineering Identity,” *Proc. from ASEE 2016 Am. Soc. Eng. Educ. Annu. Conf. Expo.*, p. 15, 2016.

[9] J. M. Smith and J. C. Lucena, “Invisible innovators: how low-income, first-generation students use their funds of knowledge to belong in engineering,” *Eng. Stud.*, vol. 8629, no. March, pp. 1–26, 2016.

[10] D. Verdin and A. Godwin, “First in the Family: A Comparison of First-Generation and Non-First-Generation Engineering College Students,” *Front. Educ. Conf.*, pp. 1–8, 2015.

[11] S. L. Dika and M. M. D’Amico, “Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors,” *J. Res. Sci. Teach.*, vol. 53, no. 3, pp. 368–383, 2015.

[12] Z. Hazari, G. Sonnert, P. M. Sadler, and M.-C. Shanahan, “Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study,” *J. Res. Sci. Teach.*, vol. 47, no. 8, pp. 978–1003, 2010.

[13] H. Carlone and A. Johnson, “Understanding the Science Experience of Sucessful Women of Color: Science Identity as an Analytic Lens,” *J. Res. Sci. Teach.*, vol. 44, no. 8, pp. 1187–1218, 2007.

[14] H. B. Carlone and A. Johnson, “Understanding the Science Experiences of Successful Women of Color : Science Identity as an Analytic Lens,” *J. Res. Sci. Teach.*, vol. 44, no. 8, pp. 1187–1218, 2007.

[15] G. Potvin and Z. Hazari, “The Development and Measurement of Identity across the Physical Sciences,” in *Physics Educatoin Research Conference*, 2013, pp. 281–284.

[16] T. Shih and X. Fan, “Comparing response rates in e-mail and paper surveys : A meta-analysis,” *Educ. Res. Rev.*, vol. 4, pp. 26–40, 2009.

[17] U.S. Department of Education. National Center for Education Statistics, “Bridging the gap: Academic Preparation and Postsecondary Success of First-Generation Students,” Washington, D.C., 2001.

[18] T. A. Brown, *Confirmatory Factor Analysis for Applied Research*, 2nd ed. New York: The Guilford Press, 2015.

[19] R. B. Kline, *Principles and practice of structural equation modeling*. Guilford publications, 2016.

[20] R Core Team, “R: A Language and Environment for Statistical Computing.” Vienna, Austria, 2014.

[21] Y. Rosseel, “lavaan : An R Package for Structural Equation,” *J. Stat. Softw.*, vol. 48, no. 2, pp. 1–36, 2012.

[22] S. E. Gregorich, “Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework.,” *Med. Care*, vol. 44, no. 11 Suppl 3, pp. S78-94, 2006.

[23] G. W. Cheung and R. B. Rensvold, “Assessing Extreme and Acquiescence Response Sets in Cross-Cultural Research Using Structural Equations Modeling,” *J. Cross. Cult. Psychol.*, vol. 31, no. 2, pp. 187–212, 2000.