International Journal of Engineering Education Vol. 37, No. 5, pp. 1382–1397, 2021
Printed in Great Britain
0949-149X/91 $3.00+0.00

Valuing Diversity and Enacting Inclusion in Engineering (VDEIE): Validity Evidence for a New Scale*

KAREN E. RAMBO-HERNANDEZ
Teaching, Learning and Culture, and Educational Psychology, College of Education and Human Development, Texas A&M University, USA. E-mail: rambohernandez@tamu.edu

REBECCA A. ATADERO
Department of Civil and Environmental Engineering, Colorado State University, 1372 Campus Delivery, Fort Collins, CO 80523, USA. E-mail: ratadero@colostate.edu

CHRISTINA H. PAGUYO
Office of Teaching and Learning, University of Denver, 303-871-6012, 2251 E. Evans Ave, Denver, CO 80210, USA. E-mail: christina.paguyo@du.edu

MELISSA MORRIS
University of Nevada, Las Vegas, 4505 S. Maryland Pkwy Las Vegas, NV 89154, USA. E-mail: melissa.morris@unlv.edu

SEOYEON PARK
Department of Teaching, Learning, and Culture, College of Education and Human Development, Texas A&M University, Harrington Building, 540 Ross Street, College Station, TX 77843, USA. E-mail: pseoyeon5@tamu.edu

A. M. ARAMATI CASPER
Department of Civil and Environmental Engineering and Graduate Degree Program in Ecology, Colorado State University, 1372 Campus Delivery, Fort Collins, CO 80523, USA. E-mail: aramati.casper@colostate.ed

BLAINE AUSTIN PEDERSEN
Department of Educational Psychology, College of Education and Human Development, Texas A&M University, Harrington Education Center Office Tower, 4225 TAMU, 540 Ross Street, College Station, TX 77843, USA. E-mail: blaine.pedersen@tamu.edu

JEREMY SCHWARTZ, GRADUATE STUDENT
Department of Communication Sciences and Disorders, West Virginia University. Morgantown, WV 26505, USA. E-mail: jschwar2@mix.wvu.edu

ROBIN A. M. HENSEL
Fundamentals of Engineering Program, Statler College of Engineering and Mineral Resources, West Virginia University, 1306 Evansdale Drive, Morgantown, WV 26506, USA. E-mail: Robin.Hensel@mail.wvu.edu

The purpose of this paper is to detail the initial validation of a scale to assess engineering students' attitudes toward the value of diversity in engineering and their intentions to enact inclusive behaviors. In study 1, we administered the scale four times. We subjected the first administration to exploratory factor analysis (EFA), and the remaining three administrations to both confirmatory factor analysis (CFA) and tests of longitudinal measurement invariance (LMI). All tests indicated strong evidence for the internal structure of the factor structure of the survey. The four factors were: engineers should value diversity to (a) fulfill a greater purpose and (b) serve customers better; and engineers should (c) challenge discriminatory behavior and (d) promote a healthy work environment. In study 2, we again assessed the structure of the data as described in study 1 and then used the scale to assess potential differences between undergraduate students who participated in activities designed to promote diversity, equity, and inclusion (DEI) (n = 116) and those who did not (n = 137). Students in the intervention classes demonstrated a small statistically significant increase in their intention to promote a healthy team environment in reference to the comparison classes. No differences were observed between the classes on the other factors. Future directions and implications are discussed.

Keywords: measurement; diversity; engineering students; survey; instrument development

1. Introduction

In many parts of the world, engineering is a profession with a history of exclusion. In the U.S., the disparities in engineering education and careers are glaring where Black, Hispanic, Native American or Alaska Native men and women, White women, Asian women, and people with disabilities are all underrepresented as compared to their proportion of the United States (U.S.) population [1]. But the lack of representation is not just a U.S. issue. For example, in the European Union, women account for 41% of scientists and engineers [2], and in the U.K., women make up only 12% of the engineering workforce [3]. There are many explanations for the gaps in representation in engineering degree pro-

* Accepted 23 May 2021.

grams and the lower number of degrees earned by individuals from these underrepresented categories. One frequently identified concern is the culture of engineering education, particularly a lack of diversity and limited appreciation for inclusion. A variety of structural, curricular and co-curricular steps can be taken to enhance the culture of an educational organization with regard to diversity, equity, and inclusion and to provide the groundwork for graduates to carry these more inclusive attitudes into the profession. However, no psychometrically sound measure currently exists to assess students' perceptions of student attitudes toward the value of diversity in engineering nor their intentions to enact inclusive behaviors – both important indicators of culture. The purpose of studies presented here is to detail the assessment of a new scale, namely the Valuing Diversity and Enacting Inclusion in Engineering (VDEIE), designed to measure engineering undergraduate students' attitudes toward the value of diversity and intentions to enact inclusive behaviors.

Culture is a compelling explanation for underrepresentation as studies have shown issues of culture affecting many populations of underrepresented students. In their signature book about why undergraduates leave the science, technology, engineering and mathematics (STEM) disciplines, Seymour and Hewitt [4] analyzed data from hundreds of hours of ethnographic interviews and focus groups to highlight the deleterious impacts of teaching STEM content aimed toward teaching the dominant population: white cis-gender men. Teaching toward and focusing the curriculum on the dominant culture has many negative consequences. For example, undergraduate women cite informal interactions and sexism in teams as propagating a culture that is unwelcoming to women [5]. Further, women of color are particularly negatively impacted by such culture. A recent systematic synthesis examined the social pain experienced by women of color and the navigational strategies they use to address that pain [6]. Employing strategies to address the social pain diverts their cognitive resources away from their education to managing their environment – thus creating greater inequities. Queer students, those who identify beyond cisgender and/or heterosexual binaries, experience similar unwelcoming environments that require additional navigational strategies to participate in and persist in STEM [7–9]. Additionally, students with physical disabilities have encountered not only physical barriers to participation in science and engineering laboratories, but also social barriers, such as lack of understanding from instructors, unaccommodating faculty, and assignment to observing and notetaking roles [10].

Although higher education institutions and political bodies laud the benefits of diversity [11], the presence of diversity in and of itself is insufficient toward creating robust outcomes in classrooms, teams, workplaces, and societies [12]. In addition to creating teams that represent diversity in terms of gender, race, and problem-solving perspectives [13, 14], such heterogeneous teams must be sustained through purposeful activities where people understand how diversity can help engineers make progress toward project goals [15]. Otherwise, engineers risk tokenizing and not fully engaging with diversity. Efforts aimed at broadening the participation of those who have been persistently underrepresented in engineering must address culture and attitudes toward the value of diversity and inclusive behaviors in engineering and should be part of the explicit curriculum for undergraduate students.

2. Inclusive Professional Identities

Our larger project is aimed at the development of inclusive professional identities [16], in which we apply the theoretical framework of professional identity development [17, 18]. Engineering identity development requires students to understand and define for themselves what it means to be an engineer and to negotiate their understanding of engineering with their own social identities [17–19]. As defined in Paguyo et al. [20], engineers who possess inclusive professional identities demonstrate exceptional technical skills, recognize and disrupt stereotypes and negative biases, promote inclusive behaviors on teams, and embrace the need to serve all groups of people. As part of students' inclusive engineering identity development, we focus our efforts on promoting positive attitudes toward diversity – by illustrating the benefits of diversity both inside the profession and outside for those whom engineers serve – and inclusive behaviors in teams.

To this end, we engage students at the early stages of engineering identity development and deliberately present an inclusive vision of the engineering profession through classroom-based activities, which directly align engineering content with the broader goal of developing an inclusive engineering identity. Specifically, we partner with engineering faculty to integrate activities into their existing course curriculum to help students (a) value diverse perspectives of their teams as an asset to problem-solving, (b) consider issues of equity and how their products or services might impact people who are different from them, and (c) enact inclusive behaviors on their teams. Consistent with Page [12], in this project we define diversity quite broadly to include diversity based on social identities, such as race/ethnicity, gender, and sexual orientation, as well as individual and cognitive dimensions of diversity, such as differences in backgrounds, expertise, and experiences. Our definition of diversity is consistent with the results of a recent study of how students in the U.S. perceived diversity, which included race, country of origin, gender, engineering discipline, and approaches to solving problems [21].

Our approach is consistent with the well-established Theory of Reasoned Action [22], which states that behaviors are a function of attitudes and perceived subjective norms. Thus, contextualized in our larger study, students' likelihood to enact inclusive behaviors is a function of their attitudes toward diversity, inclusion, and equity as well as their perceived norms of how relevant groups perceive diversity, with the group norm of interest to the research team being how students perceived the campus climate toward diversity, equity, and inclusion. Existing scales [23] assess student perceptions of the campus climate toward diversity, equity, and inclusion; however, no scales were available to assess student attitudes toward the value of diversity, equity, and inclusion within the context of engineering nor how likely students were to enact inclusive behaviors on teams. To this end, here we detail the refinement and initial validation of a scale we developed [24] to assess engineering students' attitudes toward the value of diversity in engineering and their intentions to enact inclusive behaviors.

Engineering students' attitudes toward diversity must go beyond a general appreciation of diversity and be specific to the context of engineering if students are to act on those attitudes and enact inclusive behaviors in their engineering courses and professional practice. Consistent with the Theory of Reasoned Action and drawing from the literature, we sought to create items contextualized in engineering to assess student attitudes toward their value of diversity and their intended behaviors.

2.1 Attitudes toward the Value of Diversity in Engineering

One reason to value diversity in engineering is to address social justice concerns. Social justice can be defined as ". . . full and equal participation of all groups in a society that is mutually shaped to meet their needs. Social justice includes a vision of society that is equitable, and all members are physically and psychologically safe and secure" [25, p. 1]. This definition relates to engineering in more than one way. First, to achieve social justice, all members of society with the interest and aptitude must have the opportunity and support to fully participate in engineering practices that simultaneously shape technology to meet their needs and benefit from the economic opportunity associated with engineering careers [26]. Second, as described in the preamble to the National Society of Professional (NSPE) Engineers Code of Ethics, "Engineering has a direct and vital impact on the quality of life for all people" [27], and thus engineers must create products and design solutions to meet the needs of all people. Numerous historical examples demonstrate how engineering/scientific/technical teams have contributed to biased and inequitable designs and products, such as automobile testing standards based solely on an average male occupant [28], the destruction of neighborhoods and communities of color to build the interstate highway system [29], and higher rates of funding for coronary artery disease for men despite women being at higher risk [30]. Thus, the social justice perspective values DEI because of its foundation in taking humane and moral actions [31].

Another pragmatic and instrumental rationale behind many diversity efforts in the engineering context is to improve the bottom line. A recent report studied the financial performance of companies in the UK and North America and found that companies with leadership demographics "in the top quartile of racial/ethnic diversity were 30 percent more likely to have financial returns above their national industry median" [32, p. 1] while companies in the bottom quartile for both ethnic/racial and gender diversity lagged behind in their industry. While the demonstrated link between diversity and financial performance is not causal, several possible reasons behind the link have been hypothesized, such as the ability of diverse engineers to better understand customer needs and to design improved products. This link is particularly relevant considering the changing demographics of the United States [33].

Additionally, diversity can be used as a mechanism to improve the work environment. Because of the power of diversity to foster creativity and provide new perspectives on a problem, diverse teams are more capable of solving truly difficult problems than teams of similar "smart" people [12]. Diversity in leadership has also been shown as a way of attracting and retaining the best talent [32]. For example, women have been shown to leave the engineering profession due to poor workplace climate [34].

2.2 Inclusive Behaviors in Engineering

While student attitudes toward diversity are an important characteristic, we argue that behaving inclusively is an equally, if not more, critical trait that educators can teach students as they develop inclusive professional identities. For example, engineering students should learn how to enact inclusion by valuing all team members, creating an environment free of discrimination and bias in teams, and leveraging diversity to improve teams.

This notion of teamwork as critical to engineering work is supported by statements from the National Academy of Engineering [35] and ABET, the engineering accreditation body [36, 37]. More specifically, NAE has articulated the important role teamwork plays in the engineering profession, and ABET accreditation criteria effective beginning with the 2019-2020 review cycle require engineering programs to show their graduates possess "an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives" [38, p. 40]. Student teams that work effectively exhibit positive interdependence, which is the relationship between individual success and team success [39], where a team cannot experience overall success unless each member on the team experiences success. Arguably, opportunities for students to engage in robust teamwork can be facilitated when educators increase students' awareness about the benefits of valuing all team members, which is often implied but not explicitly discussed in teamwork literature.

A prerequisite for collaborating productively is to purposefully design and facilitate a robust environment where people recognize and work to address their own biases. According to Cooper [40], teams function better when space and bandwidth exist for team members to reflect on how well they work together. While overt forms of discrimination and bias exist, there are implicit forms of discrimination and bias as well. When educators organize curricular and co-curricular experiences for students to reflect on their own potential biases and discriminatory actions, students can learn to acknowledge and act to address their explicit and implicit biases.

2.3 The Current Study

When we first attempted to assess the inclusive professional identity development of students [16], we quickly discovered there were no psychometrically sound measures to assess engineering students and their attitudes toward the value of diversity and their intentions to enact inclusive behaviors in the context of engineering. While some scales exist that address diversity broadly (e.g., *Appreciation of Cultural and Ethnic Diversity* scale, [41]) no scales existed that captured student attitudes toward the value of diversity within engineering specifically or how strongly students intended to enact inclusive behaviors. As a result, we created a new scale (see [24] for details on the initial creation of the scale and pilot study). Thus, the purpose of this manuscript is to present two studies to further detail the refinement and initial validation of a scale to assess engineering students' attitudes toward the value of diversity and their intentions to enact inclusive behaviors, namely the Valuing Diversity and Enacting Inclusion in Engineering scale. Specifically, our research questions for study 1 and 2 respectively were: (a) Does the VDEIE scale accurately and consistently measure students' attitudes toward the value of diversity and intentions to enact inclusive behavior? and (b) Are the scores obtained using VDEIE sensitive enough to detect differences between students who participated in diversity promoting activities and those who did not?

As mentioned above, we conducted a pilot study and followed the recommendations of Netemeyer, Bearden, & Sharma [42] for item development and initial validation by experts. Based on the literature previously described, we developed items to capture student attitudes toward the value of diversity, equity, and inclusion and student intentions to enact inclusive behavior on teams. According to the theory of reasoned action, intentions are reliable indicators of actual behavior [43], although the strength may be culturally dependent. The results from this first pilot administration were subjected to exploratory factor analysis [24], and the results indicated students responded to the items such that four factors were evident: *fulfill* a greater purpose, *serve* customers better, *promote* healthy behavior versus *challenge* discriminatory behavior.

2.3.1 Research Team Positionality

The team is collectively committed to creating spaces where engineering students, regardless of their backgrounds, are welcomed, appreciated, and respected. We seek to augment existing programs targeted at underrepresented students by attempting to change the culture of engineering to be more welcoming and supportive of all students, but particularly students of color, women, and queer students. Individually, our team espouses multiple lenses and identities, and when trying to assess the impact of curricular changes, we noted the lack of psychometrically sound instruments relevant to our key outcomes, and thus became impetus for the studies presented here. Of note, no one on the research team had direct contact with the participants.

3. Study 1: VDEIE Refinement Methods

3.1 Participants and Procedure

This validity study sits within a larger study focused on changing engineering curricular activities; our sample are all first-year engineering students who participated in the study. The study was deemed exempt by the university's Institutional Review Board and consent was collected within the survey. All students were from a large, R1 rocky mountain public university and enrolled in at least one of three first introductory engineering classes (mechanical engineering, civil and environmental engineering, and a general engineering course that covered multiple engineering disciplines). All students were invited to participate. Out of approximately 400 invitations to participate, 326 students responded to the survey (82% response rate). Students responded to the survey via an online platform. Students were mostly first-year (82%) and white (89% with 7% identifying as Hispanic). In addition, 67% of students identified as male, 32% of students identified as female, and 1% identified as a gender other than male or female (e.g., gender fluid, genderqueer, or non-binary). The students were mostly from civil (26%), environmental (16%), mechanical (34%), or open option engineering (no specific discipline selected, 11%).

Students responded to the scale five times throughout the course of the semester. The first administration of the survey was used in the pilot study previously mentioned [24]. The second response set was subjected to an exploratory factor analysis detailed below. The final three response sets were subjected to confirmatory factor analysis and used to determine whether the factors demonstrated longitudinal measurement invariance.

3.2 Measures

To assess the two factors, fulfill a greater purpose and serve customers better, students were prompted by the following statement, "Engineers should value diversity in order to" followed by a series of statements, such as "better serve a diverse population." Students were asked to respond on a Likert scale (1-strongly disagree to 7-strongly agree). To assess the two factors challenging poor behavior and promoting healthy behavior, students were prompted by the following statement, "While working on a team, I" followed by a series of behaviors, such as "encourage every team member to share their perspective." Students then responded to each statement on a Likert scale (1-very unlikely to 7-very likely).

3.2.1 Factor Extraction and Item Retention.

We used the results from the first administration [24] to revise the survey. For the second administration described here, we trimmed items that did not meet the specified criteria in the pilot study and added items to address the serve customers better as this factor had the fewest items relative to the other factors. Consistent with the results of the pilot study, we hypothesized a four factor solution.

Because we both trimmed and added items for the second administration of the scale, we conducted an exploratory factor analysis on the data in the second administration. We applied exploratory factor analysis (EFA) with principal axis factoring [44] to the data using direct oblimin rotation. We examined the Kaiser-Guttman rule, scree plot, parallel analysis [45], and Velicer's minimum average partial [46] test to determine the number of factors. To maintain simple structure, on the initial extractions, items were retained if the item had a pattern coefficient of at least 0.40 on the primary factor and less than 0.30 on any secondary factor. All exploratory factor analyses were conducted in SPSS v. 25. Additionally, to create a short, efficient scale to assess engineering students' valuing of diversity and intention to enact inclusive behaviors, we selected only the four highest loading items on each of the factors.

3.2.2 Confirmatory Factor Analysis and Longitudinal Measurement Invariance

To assess the stability of the factor structure, we subjected the data from the third, fourth, and fifth response sets to confirmatory factor analysis. We examined the chi-square test to assess model fit. However, the chi-square test can be oversensitive with larger samples, so we used the following indicators of acceptable model fit: Comparative Fit Index (CFI) exceeding 0.90 [47], root mean square error of approximation (RMSEA) below 0.08 or a 90% confidence interval that contained 0.05 [48], and standardized root mean squared residual (SRMR) of 0.08 or less [49]. To provide initial information for discriminatory validity, we examined the correlations of the factors at each time point.

To assess longitudinal measurement invariance, we assessed the increasingly restrictive models of configural invariance, metric invariance, and scalar invariance. The more restrictive model was compared to the less restrictive model using changes in the CFI values. If the change in CFI between the more and less restrictive models was less than 0.01, the more restrictive model was retained [50]. All confirmatory factor analyses and tests of longitudinal invariance were conducted in M*Plus* v. 7.02.

4. Study 1: VDEIE Refinement Results

4.1 Exploratory Factor Analysis

For the data obtained in the second administration of the scale, the MAP, scree plot, and Kaiser-Guttman rules all indicated a four-factor solution, and

Table 1. Reliabilities and Descriptive Statistics for Valuing Diversity and Enacting Inclusion in Engineering

Broad Construct	Factor	*n*	*r*	*M*	SD
Valuing Diversity	Fulfill a Greater Purpose	267	0.90	5.82	1.07
	Serve Customers Better	267	0.81	6.05	0.83
Inclusive Behaviors	Challenge Discriminatory Behavior	266	0.89	5.50	1.19
	Promote a Healthy Work Environment	267	0.90	6.14	0.64

the parallel analysis indicated a three-factor solution. Like the pilot administration, the results indicated a four factor solution and all factors demonstrated acceptable reliability (Table 1). The extracted factors explained 78% of the variance in the data. See Table 2 for the final retained items and factor loadings. As mentioned above, the final items retained on the scale were the four highest loading items on each factor. However, we chose to retain five items on the challenge discriminatory behavior because the item with the fifth-highest loading, "challenge homophobic behavior," was a discriminatory behavior that needed to be represented in the list. Had we left it off, we would have discriminated against a group of people who needed to be represented. The four factors are described below.

4.1.1 *Fulfill a Greater Purpose*

A high score on this factor indicated the engineering student perceived valuing diversity aligned with a strong inward desire for purpose and fairness in their work. As shown in Table 1, students mostly agreed with the statements with a moderate amount of variation.

4.1.2 *Serve Customers Better*

A high score on this factor indicated the engineering student believed customers are better served when diversity is valued. The mean scores of students, illustrated in Table 1, indicate that students strongly endorsed the importance of diversity in service to customers. Also, the smaller deviation, in comparison to fulfilling a purpose, shows that students generally agreed valuing diversity promoted better customer service.

4.1.3 *Challenge Discriminatory Behavior*

A high score on this factor indicated that the engineering student would call out any type of discriminatory behavior while working on a team.

Table 2. Pattern Matrix with Factor Loading for the Final Items

Pattern Matrix					
Engineers should value diversity to:		**Fulfill**	**Serve**	**Challenge**	**Promote**
F1	Fulfill a social responsibility for making the world better	**0.87**		0.11	
F2	Work for a greater cause	**0.84**	–0.12		
F3	Help improve the bottom line	**0.82**	–0.13		
F4	Do the right thing	**0.95**			
S1	Help them understand client and customer needs		**–0.81**		
S2	Improve products		**–0.93**		
S3	Increase public access to technology and engineered products	0.11	**–0.84**		
S4	Collaborate effectively with stakeholders in an engineering project	0.26	**–0.69**		
While working on a team, I:					
C1	Challenge homophobic behaviors		–0.11	**0.75**	–0.13
C2	Challenge racist behaviors			**0.94**	
C3	Challenge any type of discriminatory behaviors			**0.93**	
C4	Challenge sexist behaviors			**0.87**	0.12
C5	Challenge xenophobic behaviors, which are behaviors that discriminate against people from other countries			**0.82**	0.11
P1	Include everyone in all team meetings				**0.89**
P2	Make sure to give credit to team members who make contributions to the project	–0.17	–0.11	0.16	**0.65**
P3	Make sure all team members have the opportunity to take part in decision-making		–0.12		**0.82**
P4	Make sure every team member has the opportunity to contribute to the project	0.14	0.11		**0.86**

Note: Factor loadings less than or equal to |0.10| are not shown for simplicity.

As indicated in Table 1, compared to the other factors, scores for this factor were the lowest with the largest standard deviation. In general, students were only somewhat likely to intend to behave in ways that challenge discriminatory behaviors.

4.1.4 Promote a Healthy Team Environment

A high score on this factor indicated the engineering student would take measures to ensure every team member was included and valued and sought to have a variety of skills represented on the team. In contrast with challenging discriminatory behavior, students strongly agreed with the statements in promoting healthy team environments, as shown in Table 1. This relatively high mean and small standard deviation show students more readily endorsed promoting a healthy team environment than challenging discriminatory behaviors.

All of the factor means were positively correlated. The fulfill a greater purpose reason to value diversity factor was positively correlated with serve customers better (r = 0.66), challenge discriminatory behavior (r = 0.33), and promote a healthy team environment (r = 0.36). Serve customers better was positively correlated with challenge discriminatory behavior (r = 0.43) and promote a healthy team environment (r = 0.51). Finally, challenge discriminatory behavior and promote a healthy team environment were also positively correlated (r = 0.51).

4.2 Confirmatory Factor Analysis

After the pilot and second administration, we administered the scale three times over the course of approximately two months to the same group of students previously described. With the data collected in each administration, we conducted three separate CFAs for the four-factor model. The three models were named Time 3, Time 4, and Time 5, respectively (See Table A1 supplementary materials. The Chi-square values for Time3 (χ^2 = 307.04, df = 113, p < 0.001), Time4 (χ^2 = 376.75, df = 113, p < 0.001), and Time5 (χ^2 = 250.03, df = 113, p < 0.001) were statistically significant, which is a common result when using the Chi-square statistic with large samples. The CFI values suggest that all three models obtained values of >0.90, which is regarded as acceptable models. Time 5 presented a particularly good fit with a CFI value of 0.958. The RMSEA value of Time 5 also exhibited adequate fit (<0.08) while the Time 3 and 4 models did not. Nevertheless, the SRMR values of the three models showed a good fit (<0.08). In general, the results indicated that the four-factor model had a reasonable goodness-of-fit overall. Furthermore, all items exhibited high factor loadings throughout the three factor models, see Table 4. The Cronbach's alpha and composite reliability (CR) of the Time 5 model were calculated for the internal consistency reliability measures (Table 3). All factors were found to be internally consistent considering both Cronbach's alpha and CR values exceeded 0.70 [47]. The correlations between factors (see supplemental materials Tables A2, A3, A4) were moderately stable across three models except for the correlation between "Fulfill a greater purpose" and "Serve customers better" were higher at times 4 and 5.

4.3 Longitudinal Measurement Invariance

The results of the longitudinal measurement invariance are presented in Table 4. Without longitudinal measurement invariance, we cannot reliably

Table 3. Reliability measures for the factors

		Factor loading			Reliability Time 5	
Factor	**Item**	**Time3**	**Time4**	**Time5**	**CR**	**Cronbach's alpha**
Fulfill a greater purpose	F1	0.919	0.901	0.868	0.936	0.9366
	F2	0.916	0.936	0.925		
	F3	0.818	0.868	0.873		
	F4	0.875	0.769	0.879		
Serve customers better	S1	0.799	0.841	0.886	0.933	0.9306
	S2	0.850	0.897	0.932		
	S3	0.839	0.904	0.881		
	S4	0.774	0.782	0.823		
Challenge discriminatory behavior	C1	0.605	0.751	0.678	0.935	0.9192
	C2	0.902	0.948	0.934		
	C3	0.956	0.964	0.905		
	C4	0.918	0.882	0.903		
	C5	0.828	0.840	0.866		
Promote a healthy team environment	P1	0.722	0.826	0.884	0.909	0.9070
	P2	0.835	0.860	0.888		
	P3	0.659	0.757	0.783		
	P4	0.831	0.794	0.822		

* CR: composite reliability.

Table 4. Study 1: Model fit indices of nested longitudinal invariance models

	Model	$\chi 2$ (df)	Model Comparison	RMSEA	RMSE 90% CI	CFI	Δ CFI	SRMR
Fulfill a Greater Purpose								
F1	Configural	81.74 (39)***	n/a	0.065	0.045, 0.084	0.98	n/a	0.027
F2	Metric	103.66 (45)***	F1	0.071	0.053, 0.088	0.975	0.005	0.074
F3	Scalar	117.19 (53)***	F2	0.068	0.051, 0.085	0.972	0.003	0.083
Serve Customers Better								
S1	Configural	116.26 (39)***	n/a	0.087	0.069, 0.105	0.961	n/a	0.044
S2	Metric	124.84 (45)***	S1	0.082	0.065, 0.100	0.96	0.001	0.074
S3	Scalar	140.41 (53)***	S2	0.079	0.064, 0.095	0.956	0.004	0.085
Challenge Discriminatory Behavior								
C1	Configural	211.92 (72)***	n/a	0.086	0.073, 0.100	0.958	n/a	0.068
C2	Metric	233.51 (80)***	C1	0.086	0.073, 0.098	0.954	0.004	0.077
C3	Scalar	257.96 (90)***	C2	0.084	0.072, 0.097	0.949	0.005	0.083
Promote a Healthy Team Environment								
P1	Configural	63.91 (39)***	n/a	0.049	0.026, 0.071	0.985	n/a	0.046
P2	Metric	72.17(45)***	P1	0.048	0.026, 0.068	0.983	0.002	0.079
P3	Scalar	89.16 (53)***	P2	0.051	0.032, 0.069	0.978	0.005	0.095

*** $p < 0.001$.

make any comparisons of the students' responses across time as the students would not be perceiving the constructs in the same way over time. Thus, longitudinal measurement invariance must first be established before proceeding with any further analyses.

The configural invariance (unconstrained) models of each factor exhibited adequate fit. This result indicates that these factors represent the data well across all times of measurement. The test of metric invariance (weak invariance) showed that the metric invariance model did not differ from the configural model (fully unconstrained) across all four factors considering that the decrease in CFI was ≤ 0.01. The scalar invariance models (full strong invariance) also did not significantly worsen CFI values compared to the metric invariance model. Therefore, the factors are concluded to have full scalar invariance over time.

5. Study 2: VDEIE Sensitivity to Intervention Methods

5.1 Participants and Procedures

To further assess the validity of the VDEIE, in fall 2017, we administered the survey at a different university, a large R1 Mid-Atlantic land grant institution. Students in a total of eight sections of a common first-year engineering course took the survey four times throughout the semester and were taught by three instructors. Of note, unlike the participants in the first study, the first-year engineering students were taught together without regard to major. Each instructor had an equal number of intervention sections (instructor A had two sections, and instructors B and C each had one section for a total of four sections, $n = 116$) and comparison sections (instructor A had two sections, instructors B and C each had one section for a total of four sections, $n = 137$).

The students in the intervention and comparison sections largely identified as White (93% and 92% respectively) and as male (72% and 75% respectively). In the intervention and comparison sections, there were few students who identified as Hispanic (1% and 4% respectively), Asian (5% in both intervention and comparison), Black (3% and 2% respectively), or a gender other than male or female (0% in all sections). Students in the intervention sections participated in multiple activities, which are described subsequently.

The activities used in the intervention courses are more thoroughly described in Paguyo et al. [51] and Atadero et al. [16], but brief descriptions follow along with the week of the semester they occurred. In the *Dean's Talk* [52] (week 2), the dean of the college of engineering spoke to students in the course to establish egalitarian norms and highlight the importance of functioning as an engineer in a global workforce. The *Teamwork Activity* (week 2) was designed to align with the aforementioned ABET outcome. Students were required to watch a video about the importance of psychological safety in teams and complete reflection questions related to the video. Next, for the *Implicit Bias Module* (week 3), students were exposed to the idea of implicit bias by watching a video, taking an Implicit Association Test of their choice, and writing a two-page reflection essay. The students also attended a *Panel of Practicing Engineers* (week

6). The panels were deliberately composed of engineers from diverse personal, educational, and professional backgrounds. The moderator posed questions focusing on topics such as the importance of teamwork and skills engineers need beyond math and science. Students were also allowed to ask questions to the panelists. After attending the panel, students completed a homework assignment that included reflection questions. Students also completed an *Iceberg Activity* (week 9) focusing on exposing how society makes assumptions about people, either consciously or subconsciously, and how those assumptions are frequently inaccurate. This activity incorporated a campus-wide reading and discussion of *Hidden Figures*. The final activity was an *Interactive Theatre Sketch* [53] (week 9). The students watched trained actors perform a sketch in which three students (two men and one woman) are working on a team project. The team has a variety of issues that lead to dysfunctional interactions. The sketch was then performed again and students from the course were invited to stop the sketch at any time and intervene as the fourth member of the team to practice interpersonal skills and mediate conflict. After each intervention, trained facilitators led the audience in providing affirmations to the student who intervened and led a discussion about how the intervention worked. As part of the course, students were required to respond to reflection questions on all of the out-of-class experiences, including the theatre sketch.

Students (n intervention = 116, n comparison = 137) took the VDEIE Scale. For this sample, the items on the scale demonstrated acceptable reliability: (a) fulfill a greater purpose (r = 0.88) and (b) serve customers better (r = 0.91), and whether the students would (c) promote a healthy team culture (r = 0.87), and (d) challenge discriminatory behavior (r = 0.93). Students took the survey four times during the semester, approximately after the first week of class, fifth week, tenth week, and thirteenth week. All reliabilities at every time point were acceptable.

5.2 Data Analysis

Before assessing any differences between the two groups of students, we first assessed the psychometric properties of the survey responses with these students using the same CFA and LMI methods described in study 1. We did this for two reasons: (a) the initial assessment of the psychometric properties was conducted with the same students as the exploratory factor analysis, and the scale should be further validated with a new sample and (b) the assessment of any differences between the two groups in study 2 is moot if longitudinal measurement invariance is not established.

After establishing the psychometric properties of the scale with this new sample, the validity argument lies in answering two complementary questions affirmatively: (a) were the two groups indistinguishable on each of the scales at pretest? and (b) did students in the intervention sections respond differently than the students in the comparison sections to the items after participating in interventions? To address the first question, we examined the mean scores of each construct prior to intervention (i.e., pretest). We hypothesized that, prior to participating in any interventions, the students in intervention sections and comparison sections would respond similarly to the scales. Specifically, we built four sequential regression models, one for each scale that first accounted for variability due to instructor and then added intervention status to the model as a predictor. If the ΔR^2 from the model that added the intervention status was not statistically significant, then the groups were determined to be the same prior to any interventions, which provides some evidence for the validity of the scales.

To address the second validity question, we built four separate two-level hierarchical linear models, one for each scale. Times one, two, and three refer to the student responses to the scales after pretest and the onset of interventions. Student mean scores of responses to each scale at times one, two, and three (level 1) were nested within students (level 2), see equation (1).

Equation (1)

Level 1:

$$y_{ti} = \pi_{0i} + \pi_{1i}(time_{ti}) + e_{ti}$$

Level 2:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + \beta_{01}(Intervention\ Status_i) \\ &\quad + \beta_{02}(mean\ at\ pretest_i) \\ &\quad + \beta_{03}(Instructor\ Effect\ A_i) \\ &\quad + \beta_{04}(Instructor\ Effect\ B_i) + r_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11}(Intervention\ Status_i) \\ &\quad + \beta_{12}(mean\ at\ pretest_i) \\ &\quad + \beta_{13}(Instructor\ Effect\ A_i) \\ &\quad + \beta_{14}(Instructor\ Effect\ B_i) + r_{1i}\end{aligned}$$

For ease of interpretation, we will describe the above equation for promote a healthy team environment scale. In this equation, y_{ti} is the mean score of promote a healthy team environment at time t for student i and is predicted from student i's mean score at the time 1 ($\pi_{0i]}$), which is the first observation after interventions began, and the expected linear change in mean scores of promote a healthy team environment ($\pi_{1i]}$) for student i for time 2 and time 3. At level 2, there are four predictors of interest: β_{00} is the predicted mean of promote a healthy team environment at time 1 for students in the comparison group controlling for instructor effects and pretest promote a healthy team environment mean; β_{01} is the predicted difference between the intervention and comparison sections on the promote a healthy team environment mean at time 1, controlling for instructor and pretest mean; β_{10} is the linear effect of time on promote a healthy team environment for a student in the comparison section, controlling for instructor effects and pretest mean; and β_{11} is the expected change from the comparison sections on the effect of time on promote a healthy team environment for students in the intervention sections, controlling for instructor effects and pretest mean. To test our hypotheses, if β_{01} is positive and statistically significant, this would indicate intervention sections on average had higher promote a healthy team environment after interventions began compared to students in the comparison sections that did not participate; and if β_{11} is statistically significant and positive, then students in the comparison section increased their promote a healthy team environment relative to the comparison sections over the semester.

6. Study 2: VDEIE Sensitivity to Intervention Results

First, descriptive statistics were analyzed across all four factors on the VDEIE scale. The mean scores across all factors were high and relatively stable (Table 5). Notably, variability across all factors was relatively small and ranged from 0.58 to 1.32 on a seven-point scale. Generally, intervention and comparison sections display similar initial means on all factors, with the largest difference observed on challenge discriminatory behavior (difference of 0.24). However, at the end of the semester, the descriptive statistics suggest that, overall, intervention sections have slightly higher mean scores across all factors, but this time, challenge discriminatory behavior had the smallest difference between the groups (difference of 0.04).

We also conducted longitudinal CFAs with the same four-factor model. The four models are named as Pretest, Time 1, Time 2, and Time 3, respectively (see Table B1 in Supplementary Materials). According to their CFI and SRMR values, the four-factor model has a reasonable goodness-of-fit overall. However, RMSEA values did not meet the Kline [48] criteria for adequate fit (< 0.08). Considering that RMSEA values are sensitive to the sample size, these high RMSEA values might be attributed to our sample size (n = 192), which may not be sufficient. The results of the longitudinal measurement invariance (see Table B2 in Appendices) show that four factors represent the data well across all times of measurement.

Having established the adequacy of the measurement model, we ran the four sequential regression models previously described for each of the four pretest means. There were no statistically significant changes in R^2 with the addition of the intervention variable: fulfill a greater purpose ($\Delta R^2 = 0.000$, $F(1, 248) = 0.014$, $p = 0.91$, serve customers better ($\Delta R^2 = 0.000$, $F(1, 248) = 0.073$, $p = 0.78$, promote a healthy team environment ($\Delta R^2 = 0.000$, $F(1, 248) = 0.047$, $p = 0.83$, challenge discriminatory behavior ($\Delta R^2 = 0.007$, $F(1, 247) = 1.818$, $p = 0.18$. The addition of the

Table 5. Descriptive statistics of the four factors of Interest by Time and Intervention Status

		Comparison			**Intervention**		
	Time	***N***	**Mean**	***SD***	***N***	**Mean**	***SD***
Fulfill a Greater Purpose	Pretest	137	5.91	1.24	115	5.89	1.22
	1	123	5.72	1.26	96	5.91	1.32
	2	118	5.68	1.22	104	5.79	1.22
	3	123	5.89	1.11	102	6.02	1.11
Serve Customers Better	Pretest	137	6.16	0.85	115	6.12	0.88
	1	123	6.03	0.98	96	5.99	1.27
	2	118	5.99	0.94	104	6.03	1.07
	3	123	6.12	0.94	102	6.25	0.89
Challenge Discriminatory Behavior	Pretest	136	5.45	1.54	115	5.69	1.20
	1	121	5.71	1.34	96	5.85	1.12
	2	116	5.61	1.40	104	5.80	1.14
	3	122	5.78	1.31	101	5.82	1.03
Promote a Healthy Team Culture	Pretest	137	6.35	0.62	115	6.33	0.70
	1	123	6.36	0.72	96	6.42	0.73
	2	118	6.26	0.92	104	6.37	0.67
	3	123	6.23	0.83	101	6.51	0.58

intervention variables did not improve model fit and provides some evidence for the validity of the scales. Prior to the students in the intervention sections participating in any interventions, student responses were not different from students in the comparison sections.

Next, after the onset of interventions at time 1, there were no differences observed across any of the factors between the intervention and comparison sections (Table 6). As the semester progressed, there was no difference in how students changed relative to the intervention sections for three of the four factors, namely fulfill a greater purpose, serve customers better, and challenging discriminatory behaviors. However, students in the intervention section increased in their intentions to promote a healthy team environment relative to students in the comparison sections (See β_{11} for promote a healthy team in Table 6). This increase is illustrated in Fig.1. While the increase is small, the variability on the scale was also small. Specifically, prior to adding any predictors to the model, the standard deviation of the slope was 0.23 (τ_{11} = 0.055), thus the increase in the slope of 0.14 for the intervention section

Table 6. Fixed and random effects of the models predicting each factor at times 1, 2, and 3

Fixed Effects	**Fulfill a Greater Purpose**	**Serve Customer Better**	**Promote a Healthy Team**	**Challenge Discriminatory Behaviors**
	Coefficient (SE)	**Coefficient (SE)**	**Coefficient (SE)**	**Coefficient (SE)**
Intercept, β_{00}	5.66 (0.13)**	5.95 (0.13)***	6.35 (0.09)***	5.71 (0.13)***
Intervention, β_{01}	0.17 (0.12)	–0.05 (0.12)	0.01 (0.08)	–0.01 (0.12)
Pretest, β_{02}	0.70 (0.05)***	0.68 (0.07)***	0.63 (0.06)***	0.66 (0.04)***
Instructor Effect A, β_{03}	0.04 (0.14)	0.05 (0.14)	–0.02 (0.09)	–0.06 (0.14)
Instructor Effect B, β_{04}	–0.1 (0.18)	0.07 (0.18)	0.01 (0.12)	0.02 (0.18)
Slope Intercept, β_{10}	0.1 (0.07)	0.01 (0.07)	–0.01 (0.06)	0.16 (0.08)
Intervention, β_{11}	–0.04 (0.07)	0.09 (0.06)	0.14 (0.05)*	–0.02 (0.07)
Pretest, β_{12}	–0.06 (0.03)*	–0.06 (0.04)	–0.10 (0.04)*	–0.05 (0.03)*
Instructor Effect A, β_{13}	–0.05 (0.08)	0.06 (0.07)	–0.08 (0.06)	–0.11 (0.08)
Instructor Effect B, β_{14}	0.10 (0.1)	0.04 (0.09)	–0.06 (0.08)	–0.16 (0.11)
Random Effects	Variance	Variance	Variance	Variance
Intercept, τ_{00}	0.41***	0.53***	0.09***	0.69***
Slope, τ_{11}	0.01	0.03*	0.01	0.28***
level-1, σ^2	0.43	0.33	0.30	0.60

Notes: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Each pretest variable was grand mean centered.

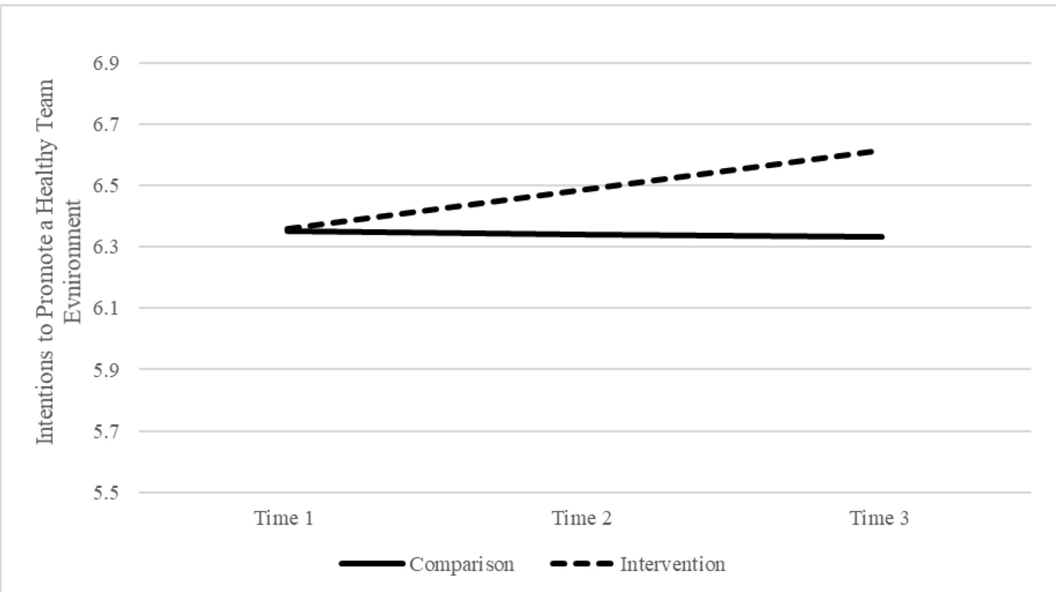

Fig. 1. Model implied intentions to promote a healthy team culture across time for engineering students in intervention and comparison classes. Note the y axis is smaller than the scale used as the standard deviation on this scale was small.

represents approximately a 0.60 change in standard deviation units. Thus, providing some validity evidence for the scale but only the factor related to student intentions to promote a healthy team environment.

7. Discussion

This study revealed several findings of interest: (a) across both samples, students responded consistently across time to the scale, providing strong evidence for the internal factor structure of this survey, and their responses were generally high and similar, (b) students were more likely to intend to promote healthy team environment than to challenge discriminatory behaviors, and (c) students who participated in interventions addressing diversity, equity, and inclusion were more likely to report higher intentions to promote a healthy team environment compared to students who did not participate in the interventions.

In both study 1 and 2 the scale was stable over time, meeting the strictest criteria of scalar invariance for each factor. Thus, the factors showed strong consistency and reliability over time. Longitudinal measurement invariance ensures the constructs being measured at each time point are the same and is a necessary condition to proceed with any further analyses. Of note, in some cases, student responses across factors started to become even more similar over time. For example, in study 1, as evidenced by the smaller correlation between the factors, students perceived differences earlier in the semester between the two factors to assess attitudes toward diversity. However, as evidenced by the larger correlation at the end of the semester, students tended to respond to the items on each factor extremely similarly by the end of the semester. This blurring of the factors at the end of the semester could indicate multiple things. One could be students did not see separate reasons to value diversity by the end of the semester – maybe these two factors are truly one factor – or the students simply experienced survey fatigue. This is a question to explore in future studies.

While there were some deflections from this trend, students generally responded on the higher end of each factor. All of the means at every time point were above neutral. These results are encouraging. At least in response to the items on the survey, students demonstrated positive attitudes toward diversity and intended to enact inclusive behaviors on engineering teams. However, one consistent difference was students indicated they were not as likely to challenge discriminatory behaviors as they were to promote a healthy team environment.

Focusing on the results of the second study, prior to any direct DEI interventions, the results indicate students in the intervention and comparison sections did not differ at pretest across any of the factors. This lack of a difference prior to interventions is a positive indication for the validity of the survey, as we would not expect students in the intervention and comparison sections to respond differently before participating in any interventions. Further, as the semester progressed, students in the intervention sections and comparison sections did not show any distinguishable difference on three of the four factors. Students in the intervention only increased relative to their non-intervention peers on one factor: their intention to enact behaviors that promote a healthy team culture. Next, we discuss some potential reasons for only observing increases on one of the four factors.

Given the nature of the interventions, lessons on promoting a healthy team environment were more overt. For example, the dean's talk established egalitarian norms, the panel discussed leveraging diversity to improve teamwork and designs, and the teamwork module directly addressed improving psychological safety by promoting healthy proactive behaviors. Upon reflection, the activities may not have equally addressed all four factors on the scale but may have focused on promoting a healthy team environment over the other three factors. The only intervention that directly addressed challenging behaviors was the theatre sketch. But even in this intervention, we anecdotally noted engineering students tended to intervene in the scene by redirecting others back to the task at hand and rarely chose to intervene by calling out the aggressive, misogynistic behavior of one of the actors. Further, we also suspect it is easier for instructors to teach how to be a better teammate than how to deal with microaggressions, racism, sexism, xenophobia, and the myriad of interactions that can occur on teams. Despite the interventions yielding some effects, more work needs to be done to make inclusion, particularly in the face of differences, more explicit.

8. Limitations

We suspect the student responses to the scale may be subject to social desirability responses bias [54, 55] or the Hawthorne effect [56]. If students were influenced by social desirability, then students may have responded positively simply because they wanted to think of themselves as having positive attitudes toward diversity and being inclusive rather than responding to the items with their true underlying attitudes. As mentioned above, the scale has only shown sensitivity to assess changes in curriculum focused on promoting a healthy team environment. The scale should be further validated by assessing the impact of activities that more directly target appreciation for diversity and preparing students to challenge discriminatory behavior.

Further, in this study, there were not enough students from traditionally underrepresented groups to explore the potential differences between the well represented and underrepresented groups. Both of the student groups in the two studies were largely White, cis-gender men first-year students at large universities, this scale should be further assessed in contexts where it is possible to disaggregate by race/ethnicity and gender identity.

In light of the recent racial reckonings in the United States and the ongoing violence against and systemic barriers facing Black, Indigenous, and other People of Color, we also suggest exploring another construct: the willingness to address systemic inequities and privilege. This potential new construct could further assess how likely engineering students are to advocate for inclusion and equity. The constructs we assessed likely do not fully capture students' intentions to elevate, promote, and advocate for diversity.

9. Conclusion

The culture in engineering has long been characterized by a lack of diversity in identities and corresponding practices that are hostile to the full participation of students who have been traditionally underrepresented in engineering. The absence of these voices of students and engineers substantially limits the positive impact engineering can have and the wicked problems engineering can solve. Given the complexity of assessing curricular change, the VDEIE scale could be used to assess student attitudes toward the value of diversity and their intentions to enact inclusive behaviors. Having a better tool to assess student attitudes toward the value of diversity and inclusive behaviors will enable researchers to take the proverbial temperature of a group of students and assess the effect of interventions developed to change the culture of engineering.

Acknowledgments – Funding provided by NSF Collaborative Grant #s: 2033129, 1725880, 1726088, 1726268, & Prior NSF funding #: 1432601. 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.

References

1. The National Science Foundation (NSF), https://www.nsf.gov/statistics/2017/nsf17310/digest/occupation/employment-sector-disability.cfm, Accessed 19 May, 2021.
2. EuroStat, https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20210210-1, Accessed 29 April, 2021.
3. L. Armitage, M. Bourne, J. Di Simone, A. Jones and S. Neave, Engineering UK 2020: Educational pathways into engineering, https://www.engineeringuk.com/research/engineering-uk-report/, Accessed 18 May, 2021.
4. E. Seymour and N. M. Hewitt, *Talking About Leaving: Why Undergraduates Leave the Sciences*, Westview Press, Boulder, CO, 1997.
5. C. Seron, S. S. Silbey, E. Cech and B. Rubineau, Persistence is cultural: Professional socialization and the reproduction of sex segregation, *Work and Occupations*, **43**(2), pp. 178–214, 2016.
6. M. Ong, N. Jaumot-Pascula and L. T. Ko, Research literature on women of color in undergraduate engineering education: A systematic thematic synthesis, *Journal of Engineering Education*, **109**(3), pp. 581–615, 2020.
7. E. A. Cech and W. R. Rothwell, LGBTQ inequality in engineering education, *Journal of Engineering Education*, **107**(4), pp. 583–610, 2018.
8. B. E. Hughes, Coming out in STEM: Factors affecting retention of sexual minority STEM students, *Science Advances*, **4**(3), p. eaao6373, 2018.
9. A. Mattheis, D. C. -R. De Arellano and J. B. Yoder, A model of queer STEM identity in the workplace, *Journal of Homosexuality*, **67**(13), pp. 1839–1863, 2020.
10. H. Jeannis, J. Joseph, M. Goldberg, K. Seelman, M. Schmeler and R. A. Cooper, Full-participation of students with physical disabilities in science and engineering laboratories, *Disability and Rehabilitation: Assistive Technology*, **13**(2), pp. 186–193, 2018.
11. C. H. Paguyo and M. S. Moses, Debating affirmative action: Politics, media, and equal opportunity in a "postracial" America, *Peabody Journal of Education*, **86**(5), pp. 553–579, 2011.
12. S. E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press, Princeton, NJ, 2007.
13. L. Hong and S. E. Page, Groups of diverse problem solvers can outperform groups of high-ability problem solvers, *Proceedings of the National Academy of Science of the United States of America*, **101**(46), pp. 16385–16389, 2004.
14. K. L. Tonso, Teams that work: Campus culture, engineering identity, and social interactions, *Journal of Engineering Education*, **95**(1), pp. 25–37, 2006.
15. C. J. Finelli, I. Bergom and V. Mesa, Student teams in the engineering classroom and beyond: Setting up students for success. Center for Research on Learning and Teaching Occasional Papers No. 29, *Center for Research on Learning and Teaching*, 2011.
16. R. A. Atadero, C. H. Paguyo, K. E. Rambo-Hernandez and H. L. Henderson, Building inclusive engineering identities: Implications for changing engineering culture. *European Journal of Engineering Education*, **43**(3), pp. 378–398, 2018.
17. M. Eliot and J. Turns, Constructing professional portfolios: Sense-making and professional identity development for engineering undergraduates, *Journal of Engineering Education*, **100**(4), pp. 630–654, 2011.
18. R. Stevens, K. O'connor, L. Garrison, A. Jocuns and D. M. Amos, Becoming an engineer: Toward a three dimensional view of engineering learning, *Journal of Engineering Education*, **97**(3), pp. 355–368, 2008.
19. F. Trede, R. Macklin and D. Bridges, Professional identity development: A review of the higher education literature, *Studies in Higher Education*, **37**(3), pp. 365–384, 2012.

20. C. H. Paguyo and T. Nosaka, Practitioner-experts and researchers in layers of learning: Researching and designing interventions in postsecondary education through the grammar of equity, in E. Mendoza, B. Kirshner and K. D. Gutiérrez (eds), *Power, Equity, and (Re)Design*, Information Age Publishing, Charlotte, NC, pp. 149–166, 2018.
21. W. C. Lee, B. D. Lutz, H. M. Matusovich and S. Bhaduri, Student perceptions of learning about diversity and its place in engineering classrooms in the united states, *International Journal of Engineering Education*, **37**(1), pp. 147–162, 2021.
22. I. Ajzen and M. Fishbein, *Understanding Attitudes and Predicting Social Behavior*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
23. G. D. Campbell-Whatley, J. A. Lee, O. Toms and C. Wang, Student perceptions of diversity and campus climate: When students speak, *New Waves-Educational Research & Development*, **15**(1), pp. 3–22, 2012.
24. K. E. Rambo-Hernandez, R. A. Atadero, C. H. Paguyo and J. Schwartz, Inclusive engineering identities: Two new surveys to assess students' inclusive values and behaviors, *American Society for Engineering Education Annual Conference and Exposition*, Columbus, OH, 2017.
25. M. Adams, L. A. Bell and P. Griffin (eds), *Teaching for Diversity and Social Justice*, 2nd edn, Routledge, New York, NY, 2007.
26. E. Torpey, Engineers: Employment, pay, and outlook, *U.S. Bureau of Labor Statistics*, 2020.
27. National Society of Professional Engineers (NSPE) Code of Ethics for Engineers, https://www.nspe.org/resources/ethics/code-ethics, Accessed 29 April, 2021.
28. K. Shaver, Female dummy makes her mark on male-dominated crash tests, *Washington Post*, 2012.
29. D. Karas, Highway to inequity: The disparate impact of the interstate highway system on poor and minority communities in American cities, *New Visions for Public Affairs*, **7**, pp. 9–21, 2015.
30. A. Holdcroft, Gender bias in research: how does it affect evidence based medicine? *Journal of the Royal Society of Medicine*, **100**(1), pp. 2–3, 2007.
31. M. S. Moses, Moral and instrumental rationales for affirmative action in five national contexts, *Educational Researcher*, **39**(3), pp. 211–228, 2010.
32. V. Hunt, D. Layton and S. Prince, Diversity matters, *McKinsey & Company*, **1**(1), pp. 15–29, 2015.
33. Change the Equation, https://files.eric.ed.gov/fulltext/ED564127.pdf, Accessed 26 April, 2021.
34. N. A. Fouad, R. Singh and M. E. Fitzpatrick, *Stemming the tide: Why women leave engineering*, University of Wisconsin-Milwaukee, WI, 2011.
35. National Academy of Engineering, *The Engineer of 2020: Visions of Engineering in the New Century*, National Academies Press, Washington, D.C., 2004.
36. The Accreditation Board for Engineering and Technology (ABET), http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2016-2017/#outcomes, Accessed 26 April, 2021.
37. The Accreditation Board for Engineering and Technology (ABET), https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2020-2021/, Accessed 3 July, 2020.
38. The Accreditation Board for Engineering and Technology (ABET), https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2018-2019/, Accessed 29 April, 2021.
39. D. W. Johnson, R. T. Johnson and K. A. Smith, The state of cooperative learning in postsecondary and professional settings, *Educational Psychology Review*, **19**(1), pp. 15–29, 2007.
40. T. Cooper, Collaboration or plagiarism? Explaining collaborative-based assignments clearly, *Essays on Teaching Excellence*, **19**(1), pp. 2007–2008, 2009.
41. C. Price, J. Williams, L. Simpson, J. Jastrzab and C. Markovitz, *National Evaluation of Youth Corps: Findings at Follow-Up*, Corporation for National and Community Service, Washington, D.C., 2011.
42. R. G. Netemeyer, W. O. Bearden and S. Sharma, *Scaling Procedures: Issues and Applications*, Sage Publications, Thousand Oaks, CA, 2003.
43. H. S. Park, Relationships among attitudes and subjective norms: Testing the theory of reasoned action across cultures, *Communication Studies*, **51**(2), pp. 162–175, 2000.
44. B. Thompson, *Exploratory and Confirmatory Factor Analysis*, American Psychological Association, Washington, D.C., 2004.
45. J. C. Hayton, D. G. Allen and V. Scarpello, Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis, *Organizational Research Methods*, **7**(2), pp. 191–205, 2004.
46. W. F. Velicer, C. A. Eaton and J. L. Fava, Construct explication through factor or components analysis: A review and evaluation of alternative procedures for determining the number of factors or components, in R. D. Goffin and E. Helmes (eds), *Problems and Solutions in Human Assessment: Honoring Douglas Jackson at Seventy*, Kluwer, Boston, MA, pp. 41–71, 2000.
47. J. Hair, R. E. Anderson, R. L. Tatham and W. C. Black, *Multivariate Data Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 2010.
48. R. B. Kline, *Principles and Practice of Structural Equations Modeling*, Guilford Press, New York, NY, 2005.
49. L. T. Hu and P. M. Bentler, Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling: A Multidisciplinary Journal*, **6**(1), pp. 1–55, 1999.
50. T. D. Little, *Longitudinal Structural Equation Modeling*, Guilford Press, New York, NY, 2013.
51. C. H. Paguyo, R. A. Atadero, K. E. Rambo-Hernandez and J. Francis, Creating inclusive environments in first-year engineering classes to support student retention and learning, *American Society for Engineering Education Annual Conference and Exposition*, Seattle, WA, 2015.
52. J. E. Bennett and D. Sekaquaptewa, Setting an egalitarian social norm in the classroom: Improving attitudes towards diversity among male engineering students, *Social Psychology of Education*, **17**(2), pp. 343–355, 2014.
53. C. Finelli and M. Kendall-Brown, Using an interactive theater sketch to improve students' perceptions about and ability to function on diverse teams, *American Society for Engineering Education Annual Conference and Exposition*, Austin, TX, June 14–17, 2009.
54. P. Grimm, Social desirability bias, in J. Sheth and N. Malhotra (eds), *Wiley International Encyclopedia of Marketing*, 2010.
55. E. Perinelli and P. Gremigni, Use of social desirability scales in clinical psychology: A systematic review, *Journal of Clinical Psychology*, **72**(6), pp. 534–551, 2016.
56. W. R. Shadish, T. D. Cook and D. T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston, MA, 2002.

Appendices

Table A1. Fit indices for the models for Study 1

Model	*n*	χ^2 (df)	RMSEA	RMSEA 90% CI	CFI	SRMR
Time 3	210	307.04 (113)***	0.09	0.078, 0.103	0.935	0.051
Time 4	190	376.75 (113)***	0.109	0.097, 0.121	0.917	0.058
Time 5	202	250.03 (113)***	0.077	0.065, 0.090	0.958	0.037

****$p < 0.001$.

Table A2. Correlation coefficients of the factors at time 3 in Study 1

Factors	1	2	3
1. Fulfill a greater purpose			
2. Sever Customers Better	0.617***		
3. Challenge discriminatory behavior	0.474***	0.377***	
4. Promote a healthy team environment	0.499***	0.381***	0.472***

Note: ****$p < 0.001$.

Table A3. Correlation coefficients of the factors at time 4 in Study 1

Factors	1	2	3
1. Fulfill a greater purpose			
2. Serve customers better	0.924***		
3. Challenge discriminatory behavior	0.544***	0.522***	
4. Promote a healthy team environment	0.515***	0.488***	0.575***

Note: ****$p < 0.001$.

Table A4. Correlation coefficients of the factors at time 5

Factors	1	2	3
1. Fulfill a greater purpose			
2. Serve customers better	0.914***		
3. Challenge discriminatory behavior	0.567***	0.552***	
4. Promote a healthy team environment	0.474***	0.523***	0.618***

Note: ****$p < 0.001$.

Table B1. Fit indices for the models in study 2

Model	n	χ2 (df)	RMSEA	RMSEA 90% CI	CFI	SRMR
Time 1	192	316.97 (113)***	0.097	0.084, 0.110	0.892	0.061
Time 2	192	303.83 (113)***	0.094	0.081, 0.107	0.908	0.049
Time 3	192	306.04 (113)***	0.094	0.082, 0.107	0.91	0.053
Time 4	192	262.939 (113)***	0.083	0.070, 0.096	0.924	0.059

****$p < 0.001$.

Table B2. Model fit indices of nested longitudinal invariance models for Study 2 1. Fit indices for the models for Study 1

Model		χ2 (df)	Model Comparison	RMSEA	RMSE 90% CI	CFI	Δ CFI	SRMR
Fulfill a greater purpose								
F1	Configural	179.409 (74)***	n/a	0.086	0.070, 0.102	0.953	n/a	0.044
F2	Metric	198.162 (83)***	F1	0.085	0.070, 0.100	0.949	0.004	0.072
F3	Scalar	227.079 (94)***	F2	0.086	0.072, 0.100	0.941	0.008	0.079
Serve Customers Better								
S1	Configural	170.943 (74)***	n/a	0.083	0.066, 0.099	0.95	n/a	0.062
S2	Metric	200.293 (83)***	S1	0.086	0.071, 0.101	0.94	0.01	0.099
S3	Scalar	215.868 (94)***	S2	0.082	0.068, 0.097	0.937	0.003	0.103
Challenge Discriminatory Behavior								
C1	Configural	151.813 (74)***	n/a	0.074	0.057, 0.091	0.936	n/a	0.05
C2	Metric	161.380 (83)***	C1	0.07	0.054, 0.086	0.935	0.001	0.076
C3	Scalar	179.231 (94)***	C2	0.069	0.053, 0.084	0.929	0.006	0.084
Promote a Healthy Team Environment								
P1	Configural	316.429 (134)***	n/a	0.084	0.072, 0.096	0.944	n/a	0.043
P2	Metric	331.128(146)***	P1	0.081	0.070, 0.093	0.944	0	0.052
P3	Scalar	358.284 (160)***	P2	0.08	0.069, 0.091	0.94	0.004	0.055

****$p < 0.001$.

Karen E. Rambo-Hernandez, PhD is an associate professor at Texas A&M University in the College of Education and Human Development with a joint appointment in the Department of Teaching Learning and Culture and the Department of Educational Psychology. Her research interests include novel applications of multilevel modeling and growth modeling, the assessment of educational interventions to improve STEM education, and access for all students – particularly high achieving and underrepresented students – to high quality education.

Rebecca A. Atadero, PhD is an associate professor in the Department of Civil and Environmental Engineering at Colorado State University specializing in structural engineering and engineering education. Her research interests include diversity, equity, inclusion and justice in engineering education and practice, infrastructure asset management, and the intersections between engineering and public policy.

Christina H. Paguyo, PhD is the director of academic assessment in the Office of Teaching and Learning at the University of Denver. Her research interests grow from long-standing commitments to learning sciences, critical theories, and equity in education policy and praxis.

Melissa Morris, PhD is an assistant professor in residence in the Mechanical Engineering Department at the University of Nevada, Las Vegas. Her research interests include retention and support of traditionally underrepresented students in STEM fields and the assessment of curricular modifications in engineering courses.

Seoyeon Park, PhD is a researcher in the Department of Teaching, Learning, and Culture at Texas A&M University. Her research interests include technology-supported solutions for struggling/underrepresented students in STEM fields.

Aramati Casper, PhD is a research scientist at Colorado State University. Her research interests center on social justice and equity in STEM, and span the natural and social sciences. She is interested in how to create systemic change in undergraduate STEM education and STEM practices.

Blaine Austin Pedersen is a graduate student at Texas A&M University in the College of Education and Human Development, Department of Educational Psychology. His research interests include the application of covariance structure analysis to study the psychosocial factors that influence the retention of students, particularly those who are traditionally underrepresented, in STEM fields and the evaluation of interventions in educational settings to advance STEM education.

Jeremy Schwartz, AuD is a recent graduate of the Department of Communication Sciences and Disorders at West Virginia University. He served as a research assistant with Dr. Rambo-Hernandez and is now working as an audiologist in North Carolina.

Robin A. M. Hensel, EdD is the Assistant Dean for Freshman Experience in the Fundamentals of Engineering Program of the Benjamin M. Statler College of Engineering and Mineral Resources at West Virginia University. Her research interests include the assessment of educational interventions and innovations to improve STEM education at all levels, facilitate high school to college transition, and increase student retention, persistence, and degree completion in STEM fields.