BOARD # 313: [NSF:EEC CAREER]Fair Assessment: Year Three of Examining Cultural Familiarity to Decrease Bias in Engineering Classroom Assessments

Dr. Kerrie A Douglas, Purdue University at West Lafayette (PWL) (COE)

Dr. Douglas is an Associate Professor in the Purdue School of Engineering Education. Her research is focused on improving methods of assessment in engineering learning environments and supporting engineering students.

Ms. Tiantian Li, Purdue University at West Lafayette (COE)

Tiantian Li (Olivia) is a dedicated Ph.D. student in Engineering Education at Purdue University. She received her Bachelor's degree in Biological Engineering, with a concentration in Pharmaceutical Processing Engineering. Currently, she is pursuing her Master's degree in Industrial Engineering. Tiantian has significant research experience in instrument development and validation analysis. Her research interests lie in developing reliable and valid measures for assessing complex engineering competencies, such as systems thinking skills. More particularly, she's interested in assessing engineering students' socio-technical systems thinking skills during their design process. In addition to her work on instrument development, Tiantian is also passionate about exploring the experiences of international scholars in the United States

Shauna Adams, Purdue University at West Lafayette (COE)

Shauna is currently a PhD student in Engineering Education at Purdue University. She previously worked for 10 years as a Systems Surety Engineer and 15 years as a coordinator/advisor of various pre-college engineering programs. Shauna earned her BS in Mechanical Engineering from North Carolina Agricultural and Technical State University and her MS in Mechanical Engineering from The Ohio State University. Her research interest includes understanding the impact of community-based engineering programs on student major choice for underrepresented groups in engineering, with a specific focus on Black/African American youth and inclusive program evaluation methods. Shauna is engaged in Purdue's Engineering Education Graduate Student Association and Black Graduate Student Association.

NSF Grantee Poster Session

[NSF:EEC CAREER] Fair Assessment: Year Three of examining cultural familiarity to decrease bias in engineering classroom assessments

Introduction & Project Motivation

This poster paper presents an overview and research progress for the CAREER project funded by NSF Division of Engineering Education and Centers (EEC). The poster will present the studies conducted in this project and display the results of analyses. The project focuses on examining how cultural familiarity plays a role in racially minoritized students' experience with engineering classroom assessments. Historically, the field of engineering is confronted with the issue of underrepresentation and the continual achievement gaps between minoritized students and majority students in engineering education (e.g., between women and men students and racially minoritized and racial majority students, etc.; [1], [2], [3]). Given the amount of emphasis of engineering education practitioners place on test scores when evaluating student performance, it is necessary for researchers to examine how engineering assessments such as exams contribute to the consistent achievement gap between different groups of students. Moreover, it is vital for researchers to abandon the deficit-based mindset when interpreting student test scores, i.e. making references of the ability of succeeding in engineering among a group of students and determining their "fitness" in this field based on test scores [4]. Instead, more discourse is needed to understand the achievement gaps from the perspectives of how engineering assessments are designed and whether such design contain implicit bias towards students from certain cultural groups. In other words, researchers need to examine whether engineering assessments are created to fairly measure what students from all cultural backgrounds know about the engineering subjects of interest. Unless more efforts are created to examine engineering assessment fairness, the continual achievement gap will continue to contribute to the underrepresentation in the engineering profession while preventing many talented students from pursuing an engineering career.

Thus, the overarching research question for this project is centered around: *why are there score disparities between students from certain cultural backgrounds, more specifically, between racially minoritized students and racial majority students?*

Project Overview and Methods

This project adopts an exploratory sequential mixed-method design that consists of a quantitative phase and a qualitative phase [5]. Table 1 illustrates the research phases and subsequent research questions in each phase. Additionally, this project specifically examines three widely used concept inventories (CI) in engineering classrooms and how the CI items function for racially minoritized and racial majority students. In the quantitative phase, we examine the item functioning of the CI items using the Classical Test Theory (CTT) as a robust alternative to more advanced psychometric item-functioning analyses such as the Item Response Theory (IRT), which requires large group size and is challenging to achieve in engineering educational research due to underrepresentation [6]. In the quantitative phase, items with problematic functioning in terms of difficulty and discrimination for racially minoritized students are identified. The findings of the quantitative phase inform the research design and result interpretation of the following qualitative phase.

The qualitative phase of this project consists of two stages. In the first stage, we analyzed the contexts of the problematic CI items identified in the quantitative phase. The purpose of context analysis is to examine the item design from the lens of cultural familiarity and reveal potential underlying bias that may disadvantage students from certain culture groups, races, or ethnicities. In the second phase, we conducted individual semi-structured interviews with racially minoritized engineering students. The purpose of the interviews are two-folds: 1) to understand their experiences taking engineering classroom exams, and 2) to conduct think-aloud interviews with the problematic CI items identified in the quantitative phase and reveal students' cognitive patterns. Ultimately, the interviews will generate insights on what students find problematic about engineering assessments and their design. Increased understanding of students' experiences with exams can inform more effective and culturally inclusive assessment practices.

Table 1 *Project timeline and design overview*

Phase	Method	Subsequent Research Questions	Data Collection Method
1	Quantitative	To what extent do items from commonly used engineering CIs demonstrate acceptable functioning (in terms of difficulty and discrimination) for racially minoritized students when compared to racial majorities?	Student responses to commonly used CIs from existing database.
2	Qualitative	What are the patterns of cultural familiarity and content of problematic CI items and CI items that show acceptable functioning?	Open-ended coding of CI item contexts from the lens of cultural familiarity. Think-aloud interviews.
3	Qualitative	How do racially minoritized students experience testing in engineering classrooms?	Semi-structured interviews.

Research Progress & Results Summary

This project was funded in 2021 and is currently in Year 3. The researchers have completed the data analyses of Phase 1 and the open-ended coding of CI item contexts in Phase 2. Currently, data collection for the semi-structured and think-aloud interviews are ongoing (eight interviews from a higher education institution collected). Tables 2 and 3 provide an overview of our Phase 1 findings using data collected by instructors across multiple U.S. higher education institutions. In Phase 1, we computed item discrimination and difficulty for CI items for two commonly used engineering CI—Dynamics Concept Inventory [DCI; 7] and Thermal and Transport Science Concept Inventory [TTCI; 8]. To determine whether an item is acceptable or unacceptable, we used Jorion et al.'s CI item judging criteria [9]. Using Jorion et al.'s judging criteria, an item is classified as unacceptable in three ways: overly difficulty, overly simple, or unable to distinguish students with high conceptual understanding from others (i.e., low discrimination value). It is clear from Tables 2 and 3 that some CI items exhibit unacceptable functioning for minoritized

students while few unacceptable items were identified for racial majority students. This indicates that existing validation evidence of commonly used CIs are likely established using samples that are dominated by racial majority, retaining potential problematic items for minoritized students due to smaller group sizes.

Table 2 *Number of unacceptable items in DCI for different race/ethnic groups*

	W	AA	API	H	Other
n	310	10	86	79	36
Number of overly difficult items[a].	0	4	2	3	4
Number of overly easy items[b]	3	0	2	2	2
Number of items with low discrimination values[c].	0	6	3	6	3

Note. W = White; AA = African American; API = Asian or Pacific Islander; H = Hispanic, and Other = other races or ethnicities.
[a]Item difficulty value is between [0,0.2]
[b]Item difficulty value is between [0.8,1]
[c]Item discrimination is smaller than or equal to 0.2.

Table 3 *Number of unacceptable items in TTCI for different race/ethnic groups*

	W	AA	API	AI	H/L	Multi	Other
n	518	16	120	5	24	13	55
Number of overly difficult items[a]	0	5	0	5	0	0	1
Number of overly easy items[b]	1	0	0	3	3	0	0
Number of items with low discmination values[c]	0	5	0	6	2	2	1

Note. W = White; AA = African American; API = Asian or Pacific Islander; H/L = Hispanic or Latino; Multi = multiracial, and Other = other races or ethnicities.
[a]Item difficulty value is between [0,0.2]
[b]Item difficulty value is between [0.8,1]
[c]Item discrimination is smaller than or equal to 0.2.

Our qualitative data analyses reveal interesting findings as well. We conducted context analysis of 90 items across three commonly used engineering CIs (DCI, Heat and Energy Concept Inventory [10] and Statistics Concept Inventory [11]). Our context analysis revealed that only seven items did not contain any cultural context [12], indicating the prevalence of cultural familiarity and how it may influence student response. Several major emerging themes were identified about the CI item contexts: access to technology, culturally sensitive, insider knowledge, and assumed experiences [12]. Items with elements from these themes require students to have familiarity with different components of a specific culture. Though still ongoing, we were able to generate preliminary findings from four interviews of racially minoritized students on their experiences with engineering exams. Our preliminary findings reveal that hidden curriculum can be identified in terms of how exams are designed, administered, and how instructors communicate exam expectations in engineering courses [13].

Future Works & Broader Impacts

Future works of this project will focus on interviewing more racially minoritized students from multiple institutions as student demographics and engineering educational practices differ across institutions. It is also the future goal of the awardee to provide engineering educators and researchers with a robust method of evaluating the fairness of educational assessments with small sample size using a combination of CTT, item context analysis, and think-aloud interviews. Such combination requires less training in psychometrics, making it easier for education practitioners to use. Using this method, engineering educators can evaluate the fairness of their classroom assessments and minimize implicit bias. This is especially useful in engineering education as student population typically do not allow educators to examine assessment fairness in depth using advanced psychometric analyses due to small group size for certain student groups. Engineering instructors can also use the findings from the qualitative phase of this project to inform instructional practices that can better support students to prepare for and perform in engineering exams, challenging existing hidden curriculum in engineering education. Ultimately, the awardee aims to advance knowledge on innovative ways to ensure the fairness and validity of engineering assessments that considers the perspectives of minoritized students. On a broader scale, this project complements other efforts of increasing diversity and inclusion in engineering education, contributing to fixing systemic issues in this field that hinders the success of underrepresented students.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Award No. 2047420. 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] NSF and NCSES, “Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023,” National Center for Science and Engineering Statistics, NSF 23-315, 2023. Accessed: Sep. 18, 2024. [Online]. Available: https://ncses.nsf.gov/pubs/nsf23315/report/science-and-engineering-degrees-earned#overall-s-e-degrees-earned-by-underrepresented-minorities

[2] M.-T. Wang and J. L. Degol, “Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future

Directions," *Educ Psychol Rev*, vol. 29, no. 1, pp. 119–140, Mar. 2017, doi: 10.1007/s10648-015-9355-x.
[3] R. M. Felder, G. N. Felder, M. Mauney, C. E. Hamrin, and E. J. Dietz, "A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes," *J of Engineering Edu*, vol. 84, no. 2, pp. 151–163, Apr. 1995, doi: 10.1002/j.2168-9830.1995.tb00162.x.
[4] A. Godwin, "Sitting in the Tensions: Challenging Whiteness in Quantitative Research," *Studies in Engineering Education*, vol. 1, no. 1, p. 78, Dec. 2020, doi: 10.21061/see.64.
[5] J. W. Creswell and V. L. P. Clark, *Designing and Conducting Mixed Methods Research*. SAGE Publications, 2017.
[6] R. B. Kline, "Assessing statistical aspects of test fairness with structural equation modelling," *Educational Research and Evaluation*, vol. 19, no. 2–3, pp. 204–222, Apr. 2013, doi: 10.1080/13803611.2013.767624.
[7] G. Gray, F. Costanzo, D. Evans, P. Cornwell, B. Self, and J. Lane, "The Dynamics Concept Inventory Assessment Test: A Progress Report And Some Results," in *2005 Annual Conference Proceedings*, Portland, Oregon: ASEE Conferences, Jun. 2005, p. 10.1278.1-10.1278.16. doi: 10.18260/1-2--15378.
[8] R. A. Streveler, R. L. Miller, A. I. Santiago-Román, M. A. Nelson, M. R. Geist, and B. M. Olds, "Rigorous Methodology for Concept Inventory Development: Using the 'Assessment Triangle' to Develop and Test the Thermal and Transport Science Concept Inventory (TTCI)," *International Journal of Engineering Education*, vol. 27, no. 5, pp. 968–984, 2011.
[9] N. Jorion, B. D. Gane, K. James, L. Schroeder, L. V. DiBello, and J. W. Pellegrino, "An Analytic Framework for Evaluating the Validity of Concept Inventory Claims," *Journal of Engineering Education*, vol. 104, no. 4, pp. 454–496, 2015, doi: 10.1002/jee.20104.
[10] M. Prince, M. Vigeant, and K. Nottis, "Development of the Heat and Energy Concept Inventory: Preliminary results on the prevalence and persistence of engineering students' misconceptions," *Journal of Engineering Education*, vol. 101, no. 3, pp. 412–438, 2012, doi: doi: 10.1002/j.2168-9830.2012.tb00056.x.
[11] A. Stone, K. Allen, T. R. Rhoads, T. J. Murphy, R. L. Shehab, and C. Saha, "The Statistics Concept Inventory: A pilot study," in *33rd Annual Frontiers in Education, 2003. FIE 2003.*, Nov. 2003, pp. T3D-1. doi: doi: 10.1109/FIE.2003.1263336.
[12] S. Adams, K. Douglas, and P. Goldberg, "Work In Progress: Evaluating the Cultural Context of Engineering and Engineering-Related Concept Inventory Assessment Items," presented at the 2023 ASEE Annual Conference & Exposition, Baltimore, MD, Jun. 2023. doi: 10.18260/1-2--44246.
[13] T. Li, N. Nasiru, K. Douglas, and S. Masta, "WIP: Students' experiences with engineering classroom exams," unpublished, in *ASEE Annual Conference & Exposition*, Montreal, Canada, 2025.