Streamlining the Process of Evaluating the Education and Diversity Impacts across Engineering Research Centers

Mr. Zhen Zhao, Arizona State University, Polytechnic campus

Zhen Zhao is a Ph.D student at Arizona State University in the Fulton Schools of Engineering Polytechnic School. He earned a B.S. in Computer Science and a M.S. in Software Engineering, both from XI'an Jiaotong University in China. He also received a M.S.E in Industrial Engineering from Arizona Stat University. Zhen's research interest include engineering student mentor-ship ability development, engineering research center education and diversity impact evaluation, engineering student adaptability development, and engineering graduate student attrition. Combining strength of mathematics modeling and statistics, 5+ years experience in collegiate teaching, Zhen is passionate and dedicated in preparing future engineering work force.

Dr. Adam R Carberry, Arizona State University

Dr. Adam Carberry is an associate professor at Arizona State University in the Fulton Schools of Engineering Polytechnic School. He earned a B.S. in Materials Science Engineering from Alfred University, and received his M.S. and Ph.D., both from Tufts University, in Chemistry and Engineering Education respectively. His research investigates the development of new classroom innovations, assessment techniques, and identifying new ways to empirically understand how engineering students and educators learn. He is currently the chair of the Research in Engineering Education Network (REEN) and an associate editor for the Journal of Engineering Educaiton (JEE). Prior to joining ASU he was a graduate student research assistant at the Tufts' Center for Engineering Education and Outreach.

Dr. Alison Cook-Davis, Arizona State University

Dr. Alison Cook-Davis is Assistant Director for Program Evaluation at the Arizona State University's Office of Evaluation and Educational Effectiveness (UOEEE). She has a BA in Psychology, MS in Social Psychology, MLS Legal Studies, and a Ph.D. in Experimental Social Psychology. Prior to joining UOEEE, she supported the research and program evaluation efforts of Maricopa County Adult Probation Department, coordinated and executed the research and program evaluation for a large Department of Justice Second Chance Act grant. These efforts included monitoring, assessing, and evaluating the impacts of program outcomes. Since joining the UOEEE in 2015, Dr. Cook-Davis has led research and evaluation activities for over 50 separate grant-funded programs or initiatives funded by the National Science Foundation, U.S. Department of Education, U.S. Department of State, U.S. Department of Agriculture, National Institutes of Health, and The Kern Family Foundation. These projects have focused on the evaluation of student success, outreach impacts, innovative learning techniques, and STEM-related interventions and curricula.

Dr. Jean S Larson, Arizona State University

Jean Larson, Ph.D., is the Educational Director for the NSF-funded Engineering Research Center for Biomediated and Bio-inspired Geotechnics (CBBG), and Assistant Research Professor in both the School of Sustainable Engineering and the Built Environment and the Division of Educational Leadership and Innovation at Arizona State University. She has a Ph.D. in Educational Technology, postgraduate training in Computer Systems Engineering, and many years of experience teaching and developing curriculum in various learning environments. She has taught technology integration and teacher training to undergraduate and graduate students at Arizona State University, students at the K-12 level locally and abroad, and various workshops and modules in business and industry. Dr. Larson is experienced in the application of instructional design, delivery, evaluation, and specializes in eLearning technologies for training and development. Her research focuses on the efficient and effective transfer of knowledge and learning techniques, innovative and interdisciplinary collaboration, and strengthening the bridge between K-12 learning and higher education in terms of engineering content.

Dr. Michelle Jordan , Arizona State University

Michelle Jordan is as associate professor in the Mary Lou Fulton Teachers College at Arizona State University. She also serves as the Education Director for the QESST Engineering Research Center. Michelle's program of research focuses on social interactions in collaborative learning contexts. She is particularly interested in how students navigate communication challenges as they negotiate complex engineering design projects. Her scholarship is grounded in notions of learning as a social process, influenced by complexity theories, sociocultural theories, sociolinguistics, and the learning sciences.

Wendy M. Barnard, Arizona State University

Wendy Barnard is an Assistant Research Professor and Director of the College Research and Evaluation Services Team (CREST) at Arizona State University. Dr. Barnard received her Ph.D. from the University of Wisconsin-Madison, where she focused on the impact of early education experiences and parent involvement on long-term academic achievement. Her research interests include evaluation methodology, longitudinal research design, STEM educational efforts, and the impact of professional development on teacher performance. Currently, she works on evaluation efforts for grants funded by National Science Foundation, US Department of Education, local foundation, and state grants.

Dr. Megan O'Donnell, Arizona State University

Megan O'Donnell is a Research Professional in the College Research and Evaluation Services Team (CREST). Dr. O'Donnell received her Ph.D. from Arizona State University, where she focused on risk and resiliency processes in Mexican American adolescents. Her current research and evaluation interests include evaluation methodology, mixed methods design, evaluation of engineering and STEAM education programs. Currently, she works on evaluation efforts for the US Department of Education, National Science Foundation, local foundations, and state grants.

Dr. Wilhelmina C. Savenye, Arizona State University

Dr. Wilhelmina "Willi" C. Savenye is a Professor Emeritus of Learning, Design and Technologies / Educational Technology at Arizona State University. She is a former Education Director, and currently serves as Senior Education Advisor, for the NSF Engineering Research Center for Bio-mediated and Bio-inspired Geotechnics (CBBG). She previously taught at the University of Texas at Austin and San Diego State University. She earned her M.Ed. and Ph.D. in Educational Technology from ASU, and B.A/ in Anthropology from the University of Washington. Dr. Savenye focuses on instructional design and evaluation of technology-based and online learning systems, employing both quantitative and qualitative research methodologies. She has published over 70 articles and book chapters; made over 140 conference presentations and workshops in the U.S., Europe and Asia; been awarded numerous grants, and has produced many digital learning programs. She is Editor Emeritus of the Journal of Applied Instructional Design. She has served on the editorial boards of journals including Educational Technology: Research and Development and the Quarterly Review of Distance Education, and reviews for additional journals. She served on the editorial board for the Encyclopedia of Educational Technology and has held elected leadership positions. Dr. Savenye's instructional design and evaluation work has been conducted in such diverse settings as engineering education, school districts, museums, botanical gardens, zoos, universities, corporations, and Army tank maintenance training.

Streamlining the Process of Evaluating the Educational and Diversity Impacts of Engineering Research Centers through a Common Assessment Instrument

Introduction

The National Science Foundation (NSF) has funded 74 nationwide Engineering Research Centers (ERC) since 1985. These ERCs have been leading forces in conducting advanced complex multidisciplinary research to address critical engineering challenges [1]. A core feature of ERCs is improving engineering education and diversity experiences internally within the centers and externally for the public at large. Thus, ERCs have organized education and diversity programs to enhance engineering exposure across the lifespan, encourage diversity in the STEM workforce, and connect industry with academia. These programs have included, but are not limited to yearlong research activities involving center-affiliated undergraduate students, graduate students, and postdoctoral scholars; summer research opportunities (e.g. Research Experiences for Undergraduates (REU), Research Experiences for Teachers (RET), and Young Scholars Programs (YSP); and outreach programs for K-12 students.

A grant requirement by NSF is that ERCs annually assess the impact of their education and diversity programs [2]. As part of this standard, NSF provides a set of common outcomes for all centers to design their evaluation plans and regularly evaluates whether centers are leveraging the efforts of other ERCs. However, ERCs currently undertake evaluation in relative isolation from other ERCs despite NSF's encouragement of instrument sharing [3]. The nature of this often-solitary assessment results in each center developing and using similar but also divergent evaluation tools. A consequence of these multiple ERC assessments has created redundancy in the investment of effort and resources at each center. Divergence in data collection from these redundant efforts has made cross-center comparisons difficult, if not impossible.

A consortium of three ERCs (Center for Bio-mediated and Bio-inspired Geotechnics (CBBG), Nanosystems Center for Nanotechnology-Enabled Water Treatment (NEWT), and Center for Quantum Energy and Sustainable Solar Technologies lab (QESST)) at Arizona State University banded together to form the Tri-ERC Education Consortium (TEEC). Supported directly by NSF through supplemental funding, this collaboration sought to address this need through a streamlined common instrument that can be applied across ERCs to evaluate ERCs' impact on education and diversity [4],. The consortium brought education leaders, researchers, and independent evaluators from the participating ERCs on board to work together. This joint effort used resources from each ERC to create a common tool supported by validity evidence from a larger sample of ERC participants. This paper presents the first draft of a common instrument, discusses several types of validity evidence, and presents the next steps for our work.

Background

The impact of the educational and diversity components has been studied broadly, including undergraduate student development from research experiences [3-5], teacher growth from RET experiences [6-8], undergraduate curriculum development [5, 6], engineering undergraduate student recruitment [7, 8], K-12 student impacts from ERC outreach activities [9], graduate student mentoring skill development from summer programming [10, 11], and diversity and inclusion [12, 13]. The various evaluation efforts of each of the ERC components listed reflect overall limitations in replicability across sites. First, most studies have focused on a single type of ERC population. A few exceptions have been presented via NSF solicited reports [18-20], but such reports are rare due to the associated high cost of these efforts. Second, most studies to date have been conducted within a single ERC setting, thereby restricting generalizability and power due to a small and specific sample. While we identified one study that examined three ERC summer program participant populations simultaneously [14], such studies are rare in the literature. An ERC-specific study across multiple centers with multiple participant populations can directly address these two limitations.

Theoretical framework

The instrument proposed in this study was driven from NSF guidelines [4]. NSF suggests five categories to be evaluated by all ERC education programs: understanding of the ERC, skill sets, mentorship experience, future plans, and program satisfaction [3]. NSF also requires all funded ERCs to promote a culture of diversity and inclusion as noted in the NSF ERC solicitation:

> The culture of the ERC and teams within the ERC demonstrate an environment in which all members feel valued and welcomed, creatively contribute, and gain mutual benefit from participating. Because of the ERC's attention to diversity and culture of inclusion, participation from members of groups traditionally underrepresented in engineering as well as diverse scientific and other perspectives is required [2].

The climate of inclusion was identified as an additional evaluation category included in the instrument. The proposal instrument aims to evaluate the ERC's education and diversity impact through these six common categories.

Methods

Survey Development

A four-step iterative process was used to develop this instrument (see [4] for additional details regarding this process). The consortium first reviewed existing evaluation tools previously created by each of the three participating ERCs. Questions and items were extracted and compared for similarities and differences in content and structure. The six common categories were used to group the compiled questions and items into subcategories. Items within a given subcategory were then combined and synthesized into new agreed upon items.

Questions and items associated with the category "understanding the ERC" captured the level of understanding ERC participants had regarding various aspects of an ERC (e.g., mission). Questions and items associated with the category "skill sets" aimed to grasp the impact of ERC involvement on various participants' skills (e.g., oral presentations, finding relevant literature, etc.). Questions and items associated with the category "mentorship experience" asked participants about their mentorship experiences within the ERC. Questions and items associated with the category "climate of inclusion" examined ERC participants' perception of ERC efforts designed to promote diversity and inclusion as well as their own experience of being respected, valued, and included. Questions and items associated with the category “future plan” explored ERC participants’ future career and academic plans following their participation in the ERC. Questions and items associated with the category “program satisfaction” allowed ERC participants to give feedback on their overall ERC experience.

The final step was to develop an appropriate method for presenting items for a given category. Items pertaining to understanding the ERC, skill sets, mentorship experience, and climate of inclusion were presented using a scale from 0 to 100. Items addressing future plans were presented as binary items, while program satisfaction items were presented as a combination of Likert-type, binary, and open-ended items. Additional demographic items were collected from participants, including gender, ethnicity, academic standing, parents’ education, and veteran status.

Data Collection

The survey was distributed during the summer of 2019 to REU, RET and YSP participants across all three participating ERCs and students (including undergraduate students, graduate students and postdoctoral scholars) who were engaged with yearlong research of one ERC in the past year. (Note: The timing of each ERC’s annual site visit with NSF impacted the timing for data collection of students engaged in ERC yearlong research.) All surveys were distributed online through an online survey software platform and took approximately 20 to 25 minutes to complete. A total of 126 participants completed at least one set of questions associated with the identified six categories: 77 summer program participants and 49 yearlong students. All items were presented using a slider in which participants could choose to answer with any whole number between 0 (“not at all”) and 100 (“a great deal”).

A think-aloud protocol [15, 16] was used with three REU participants during data collection to collect face validity evidence. Each participant was asked to describe their actions, thoughts, and decision making verbally as they completed the survey. The researchers and participants held a discussion following completion of the survey that allowed participants to provide their overall feedback and researchers to ask follow-up questions. This step provided descriptive data to better understand how items were being interpreted and the rationale used by participants to answer the survey questions. Researchers present during these sessions took field notes to document their observations of participant actions.

Analysis

The analysis presented in this paper focuses primarily on the four categories: understanding of the ERC, skill sets, mentorship experience, and climate of inclusion, presented to participants using a 0 to 100-point slider (restricted to whole numbers). Additional information regarding the remaining categories (future plans and program satisfaction) will be discussed in the Conclusions and Future Work section.

Exploratory Factor Analysis

The data for all four categories were not multivariate normally distributed; Exploratory factor analysis (EFA) was conducted in Mplus using robust estimator MLR [17] and Oblique rotation Geomin [18]. The cutoff of factor loadings was set at 0.4 and 0.3 for cross-loading [19]. The benchmark for Kaiser-Meyer-Olkin (KMO) value and Bartlett's test p-value are 0.8 and 0.05, respectively [19]. The number of factors to extract was determined using parallel analysis and Scree plots. EFA was conducted on the categories of understanding the ERC (KMO value = 0.852; Bartlett's test p-value < 0.05), skill sets (KMO value = 0.913; Bartlett's test p-value < 0.05), mentorship experience (KMO value = 0.891; Bartlett's test p-value < 0.05), and climate of inclusion (KMO value = 0.409; Bartlett's test p-value < 0.05). Each category was shown to demonstrate the necessary criteria for exploratory factor analysis except climate of inclusion, which revealed an unacceptably low KMO value [19].

Correlation Analysis

A two-tailed Pearson correlation analysis was conducted over all items per category to explore the pairwise correlations among the items. When two items appeared to be highly correlated (i.e., above .85) they were either combined or one of the correlated items was removed. This high correlation indicated that the two items measured similar categories from participants. A complete set of correlation tables for each category is provided in the Appendices.

Results

Understanding the ERC

Questions evaluating participants' understanding of the ERC were presented using the following item stem: "Please reflect on your understanding of the NSF-funded Engineering Research Center (ERC). Rate your present level of understanding, as well as your level of understanding prior to participating in the ERC for each of the items below." No items in this section were shown to be highly correlated with one another (see Appendix A).

A two-factor structure emerged through EFA (Table 1): 1) present understanding, and 2) prior understanding. Both factors achieved good reliability levels; Cronbach's alpha of 0.909 for present understanding and 0.907 for prior understanding.

Table 1. *Factor structure and factor loadings for understanding the ERC*

Item	Present Understanding	Prior Understanding
Presently, I understand the mission of the ERC	0.855	
Presently, I understand the primary field(s) involved in the ERC	0.889	
Presently, I understand the connection between the ERC field(s) of study and how the ERC helps people address real world issues	0.787	
Presently, I understand the engineering problems associated with the ERC field(s) of study	0.802	
Presently, I understand the career pathway(s) associated with the ERC field(s) of study	0.782	
Prior to participating in the ERC, I understood the mission of the ERC		0.527
Prior to participating in the ERC, I understood the primary field(s) involved in the ERC		0.803
Prior to participating in the ERC, I understood the connection between the ERC field(s) of study and how the ERC helps people address real world issues		0.792
Prior to participating in the ERC, I understood the engineering problems associated with the ERC field(s) of study		0.948
Prior to participating in the ERC, I understood the career pathway(s) associated with the ERC field(S) of study		0.819

Skill Sets

Items evaluating participant skill sets were presented using the following item stem: “How much has your participation in the ERC impacted the following skills.” Two highly correlated items ($r = 0.885$) were identified in this section: “finding relevant literature,” and “making connections between existing literature and research” (see Appendix B).

A two-factor structure also emerged from the EFA (Table 2): 1) communication skills, and 2) research skills. Both factors achieved good reliability levels; Cronbach’s alpha value of 0.907 for communication skills and 0.962 for research skills. Two items - “networking across universities” and “teaching and mentoring others” - also demonstrated low factor loading issues. One item, “incorporating feedback,” cross-loaded on both factors.

Table 2. *Factor structure and factor loadings for skill sets*

Item	**Communication Skills**	**Research Skills**
Oral presentation	0.793	
Writing technical reports, conference proceedings, or journal publications	0.788	
Creating visual displays such as posters or prototypes	0.714	
Networking with industry	0.442	
Networking with my peers	0.496	
Defending a position as part of a discussion	0.545	
Speaking appropriately to different audiences	0.684	
Incorporating feedback	0.426	0.456
Collaborating with others		0.534
Formulating research questions		0.568
Analyzing data (e.g., experiment results)		0.737
Solving problems		0.755
Collecting data from lab experiments, field observations, or web searching		0.875
Keeping a record of research activities		0.922
Using lab equipment		0.656
Making connections between classroom learning and research		0.690
Determining the next step in a research project		0.811
Working independently		0.853
Conducting research in an ethical and responsible manner		0.848
Providing leadership on projects		0.650
Finding relevant literature		0.769
Making connections between existing literature and research		0.787
Managing time		0.651

Mentorship Experience

Mentorship items were presented using two different item stems (Table 3). Items 1 to 6 used, "Please rate your OVERALL mentoring experience within the ERC related to..." Items 7 to 14 used, "My mentor(s) within the ERC…". No items emerged as being highly correlated for this category (see Appendix C).

A single factor model emerged through an EFA. The factor achieved a good reliability level (Cronbach's alpha = 0.981). The think-aloud data indicated that a change should be made to the item "provided me with a role model" due to confusion. This change will be made in future versions.

Table 3. *Factor structure and factor loadings for mentorship experience*

Item	**Mentorship**
advice that supports your future plans	0.827
direction on your research project	0.774
training to support independent work	0.763
feedback that is constructive	0.824
encouragement to strive for success	0.875
inspiration to pursue a career in a STEM-related field	0.822
shared knowledge and expertise	0.868
demonstrated knowledge and expertise	0.728
established my project goals	0.768
discussed my career goals	0.695
provided me with a role model	0.665
helped me when needed	0.852
displayed professionalism	0.842
challenged me to extend my abilities	0.646

Climate of Inclusion

Questions and items associated with the category "climate of inclusion" investigate: (1) how ERC participants perceived ERC's effort to promote a diverse and inclusive culture within the ERC, and (2) to what extent the participants felt valued and included while involved in the ERC. Items 1 to 13 were used to capture the ERC participants' perception, and items 14 to 22 were adopted to capture the ERC participants' experiences (Table 4).

Table 4. *Factor structure and factor loadings for climate of inclusion*

Item	Culture	Exclusion	Acceptance
implemented policies and practices that reinforce a commitment to diversity	0.878		
implemented policies and practices that reinforce a commitment to inclusion	0.873		
implemented policies and practices that reinforce a commitment to equity	0.926		
stressed the importance of learning about diversity	0.797		
stressed the importance of learning about inclusion	0.867		
stressed the importance of learning about equity	0.827		
made diversity an essential value of the ERC	0.902		
made inclusion an essential value of the ERC	0.826		
made equity an essential value of the ERC	0.847		
treated all students fairly, irrespective of their membership in diverse groups (i.e., age, race/ethnicity, religion, gender, etc.)	0.566		
created a positive climate for those who belong to diverse groups	0.868		
provided equal opportunities and access for all to participate in ERC activities	0.840		
been encouraged by the ERC to participate in diversity and inclusion events	0.511		
heard disparaging remarks made about a member or members of the ERC due to their membership in a diverse group(s)		0.975	
witnessed or have been a victim of bias within the ERC		0.962	
witnessed or have been a victim of exclusion within the ERC		0.959	
felt accepted while interacting with ERC faculty member(s)			0.627
felt accepted while interacting with ERC mentors			0.804
felt accepted while interacting with ERC staff members			0.799
felt accepted while interacting with ERC peers within a work-related situation			0.948
felt accepted while interacting with ERC peers within a non-work-related situation			0.941
felt accepted while interacting with ERC industry partners			0.735

These 22 items were also structured using two different item stems. Items 1 to 12 used "Please rate your agreement with the following statements. Overall, the members of the ERC have ..." Items 13 to 22 used "Since participating in the ERC, I have ..."

A high tendency was demonstrated during the think-aloud activity to misinterpret items pertaining to equity. We also identified 11 pairs of highly correlated items related to climate of inclusion (range of 0.852 to 0.960). This high number of highly correlated items indicated that several items designed to capture diversity or inclusion were closely related and encouraged participants to respond to these items quite similarly (see Appendix D).

An EFA was conducted despite the low KMO because of the preliminary nature of the study; a three-factor structure emerged (Table 4): 1) center culture, 2) center exclusion, and 3) center acceptance. All three factors achieved good reliability levels: Cronbach's alpha value of 0.960 for center policy, 0.920 for center exclusion, and 0.891 for center acceptance. Center exclusion consisted of three highly correlated items with extremely high factor loadings.

Discussion and Future Work

Data analyses pertaining to the categories of understanding the ERC, skill sets, mentorship experience, and climate of inclusion have provided a basis to inform revisions to the pilot instrument. Understanding the ERC produced two factors with no items being highly correlated with one another. This result suggests no changes are needed for this category.

Skill Sets revealed a two-factor structure. Two items were found to be highly correlated, which suggests the removal of the item "finding relevant literature" in subsequent versions of the survey. The one item that cross-loaded on both factors, "incorporating feedback," will be removed to avoid potential further issues of cross-loading. Two items demonstrating low factor loadings should also be revisited to determine whether or not modified versions of these items will warrant inclusion in future iterations of the survey. These decisions will be based on applicability relative to NSF directives, applicability to the broader ERC population, and item fit with emergent factors. Minor clarifications will also be made for each of the remaining items to increase consistency and clarity.

Mentorship experience produced a single factor with no items highly correlated. Think-aloud data revealed that the item "provided me with a role model" caused confusion with some participants. This item will be removed or modified in future versions of the survey.

Climate of inclusion presented a bevy of potential changes to the survey. Items pertaining to equity had the tendency to be misinterpreted based on the dual definition for equity; being fair or impartial vs. value related to something owned. All items designed to capture equity will be subsequently removed leaving the focus of these items on diversity and inclusion. Six additional items will also be removed due to high correlations. Items selected for removal will be chosen after consulting with the diversity and inclusion directors from the three partnering ERCs. Two items - "I have felt accepted while interacting with ERC peer(s) within a work-related situation" and "I have felt accepted while interacting with ERC peer(s) within a non-work related situation" - will be combined after also revealing a high correlation. Center exclusion items with high correlations and high factor loadings suggests that these items were strongly separated from other climate of inclusion items and warrant removal from the survey. Finally, two items demonstrating low factor loading - “seen that someone like me in the ERC can be successful” and “benefited from working with people from different backgrounds in the ERC” - were deemed to be of sufficient importance to revise and re-examine in future versions of the instrument.

Items developed to evaluate future plans and program satisfaction were not analyzed as part of this study due to the design of the questions as binary, open-ended, and/or Likert-type. Strategies to update these categories will include cutting down the number of open-ended questions to reduce the time it takes to complete the survey. Items removed will be considered for potential components included in complementary qualitative instruments (e.g., interviews and focus group).

Two additional changes identified during the think-aloud portion of the study will also be considered when revising the survey. The first aims to address missing data. We identified that one of the potential reasons causing missing data in this survey was participants’ assumption of the slider anchors. For instance, participants saw the slider cursor initially pointed at level “0.” Think-aloud data suggested that participants may have assumed the answer will be recorded as “0” if the cursor was left untouched. All such responses were recorded by the system as “missing.” Failing to provide any instruction or clarification about this issue led to a significant amount of missing data on certain items (around 40%), which could be a potential reason why some items have low factor loadings. Future versions of the survey will use traditional 5-point Likert-type scales to avoid this issue.

The second change identified through the think-aloud process is the potential replacement of all instances where the term “ERC” is used with the specific ERC’s name. Adding a component that highlights the specific ERC with which the individual survey taker is affiliated will help clarify what exact activities we hope to capture through this survey. Additionally, we will also reconsider the item stems used and ensure consistency with all the changes made in the revision.

We are currently revising the instrument by addressing each issue that emerged through the think-aloud data, correlation analyses, and EFA. We intend to distribute a revised version of the instrument during Spring and Summer 2020 to collect new validity evidence that will inform a finalized version of the instrument. We plan to invite more ERCs to join and test the instrument nationwide using an online surveying tool accessible to all.

We also plan to develop complementary qualitative tools to supplement the quantitative findings and address the categories that are not so easily assessed using a survey. Additional qualitative measures would allow each ERC to collect supplementary information, based on their special interests, unique setting, or from a specific participant group. These qualitative tools, including interview and observation protocols, will also allow evaluation teams to examine specific and detailed nuances of each individual program. This greater effort will enhance overall understanding of ERC educational and diversity programming and provide the necessary data to support changes throughout the ERC program.

It is important to note that the purpose of developing a common ERC evaluation instrument is not to foster ranking of programs. Our goal is to reduce siloing of ERC evaluation efforts by facilitating cooperation and collaboration across ERCs. A shared tool allows each ERC the opportunity to reflect on how well their own ERC is doing relative to other ERCs and to identify how they can improve as an ERC. This instrument will be especially useful to new ERCs who are looking to navigate the complex nature of ERC evaluation. Overall, the idea of this common instrument is intended to share best practices for the betterment of all ERCs. This mission will be further advanced through already established, regularly scheduled virtual meetings with ERC education directors and evaluators to share best practices, lessons learned, and ideas for future programming.

Conclusion

The work presented in this paper aims to streamline the process of evaluating ERCs' educational and diversity impacts. We hope to break the mold of current ongoing efforts done in isolation through a shared common tool. Such an effort is not intended to dictate how ERCs approach improvement based on the results of the survey, but rather to bring ERCs together to learn from one another. Consistency in how each ERC evaluates the effectiveness of their education and diversity programming will streamline how ERCs report to NSF during annual site visits and clarify the process for how NSF suggests improvements.

Revisions to the instrument presented will be further tested to provide new validity evidence to support the widespread use of this instrument. The revision will address all the issues that emerged through the think-aloud data, correlation analyses, and EFAs. A revised instrument will be distributed during Spring and Summer 2020 among all ERCs interested in using this new measure. Test-retest procedures will also be considered before finalizing the instrument. The finalized instrument will then be distributed to all currently funded and newly funded ERCs for full nationwide dissemination of the instrument.

Acknowledgement

This study is directly supported by the National Science Foundation through supplemental funding under NSF Cooperative Agreement No. EEC-1449501. Any opinions, findings, conclusions, or recommendations expressed in this study are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation. The authors would also like to acknowledge the support provided by the three partnering ERCs – Center for Bio-mediated and Bio-inspired Geotechnics (CBBG), Nanosystems Center for Nanotechnology-Enabled Water Treatment (NEWT), and Center for Quantum Energy and Sustainable Solar Technologies lab (QESST) – and the associated evaluation teams – University Office of Evaluation & Educational Effectiveness (UOEEE) and College Research & Evaluation Service Team (CREST).

References

[1] The National Science Foundation, "NSF Engineering Research Centers: Creating New Knowledge, Innovations and Technologies for over 30 Years," The National Science Foundation, Arlington, VA, 2015.

[2] The National Science Foundation, "Gen-4 Engineering Research Centers (ERC) Program Solicitation," The National Science Foundation, Arlington, VA,, 2019.

[3] Court, "Best Practices Manual," 14 11 2019. [Online]. Available: http://erc-assoc.org/content/chapter-4-education-programs. [Accessed 13 3 2020].

[4] Z. Zhao, A. R. Carberry, W. Barnard, A. Cook-Davis, M. Jordan, J. Larson, M. O'Donnell, and W. Savenye, "Work in Progress: Creating Common Instruments to Evaluate Education and Diversity Impacts across Three Engineering Research Centers," in *Frontiers in Education Conference*, Cincinnati, OH, 2019.

[5] D. M. Pai, R. G. Liles, C. Lambeth, P. N. Kumta,, H. S. Borovetz, S. K. Pixley and J. Sankar, "Bootstrapping a new graduate curriculum through an Engineering Research Center," in *ASEE Annual Conference & Exposition*, Vancouver, B.C. Canada, 2011.

[6] P. Hirsch, D. Kelso, B. Shwom, J. Troy and J. Wal, "Redefining Communication Education for Engineers: How the NSF/VaNTH ERC is Experimenting with a New Approach," in *ASEE Annual Conference & Exposition*, Albuquerque, NM, 2011.

[7] T. Dahlberg, T. Barnes, A. Rorrer and E. Powell, "Improving retention and graduate recruitment through immersive research experience for undergraduates," in *SIGCSE Technical Symposium on Computer Science Education*, Portland, OR, 2008.

[8] W. J. Gonzalez-Espada and D. S. LaDue, "Evaluation of the impact of the NWC REU program compared with other undergraduate research experiences," *Journal of Geoscience Education,* vol. 54, no. 5, pp. 541-549.

[9] C. M. Lambeth, B. A. McCullough and H. R. Aschenb, "Creating a pipeline into biomedical engineering," in *ASEE Annual Conference & Exposition*, Seattle, WA, 2015.

[10] Z. Zhao and A. R. Carberry, "Developing postdoctoral scholar and graduate student mentorship ability," in *Frontiers in Education Conference*, San Jose, CA, 2018.

[11] E. Dolan and D. Johnson, "Toward a Holistic View of Undergraduate Research Experiences: An Exploratory Study of Impact on Graduate/Postdoctoral Mentors," *Journal of Science Education Technology,* vol. 18, pp. 487-500, 2009.

[12] C. L. McCullough, M. Crull and D. Thomas, "Adventures in engineering: A unique program to attract under-represented groups to engineering," *IEEE Transactions on Education,* vol. 37, no. 1, pp. 3-7, 1994.

[13] M. E. Jordan, W. Wakefield, M. DeLaRosa and C. Miller, "Building youth's socio-technical engineering knowledge through engagement in a community solar energy project," in *ASEE Annual Conference & Exposition*, Tampa, FL, 2019.

[14] C. Bruchok, B. Ricca, M. E. Jordan, W. Wakefield and N. Bowers, "Relating social network structure to uncertainty and social interaction in an engineering design challenge," in *International Conference of the Learning Sciences*, London, England, 2018.

[15] J. M. Trenor, M. K. Miller and K. G. Gipson, "Utilization of a think-aloud protocol to cognitively validate a survey instrument identifying social capital resources of engineering

undergraduates,”," in *ASEE Annual Conference and Exposition*, Vancouver, B.C. Canada, 2011.

[16] K. Ericsson and H. Simon, Protocol analysis: Verbal reports as data, Cambridge, MA: MIT Press, 1993.

[17] L. K. Muthén and B. O. Muthén, Mplus User's Guide: Statistical Analysis with Latent Variables (7th ed.)., Los Angeles, CA: Muthén & Muthén, 1998-2012.

[18] M. W. Browne, "An overview of analytic rotation in exploratory factor analysis," *Multivariate Behavioral Research,* vol. 36, pp. 111-150, 2001.

[19] D. B. McCoach, R. K. Gable and J. P. Madura, Instrument Development in the Affective Domain, School and Corporate Application (3rd ed), New York, NY: Springer., 2013.

[20] S. Sheppard, S. Gilmartin, H. L. Chen, K. Donaldson, G. Lichtenstein, Ö. Eris, M. Lande and G. Toye, "Exploring the Engineering Student Experience: Findings from the Academic Pathways of People Learning Engineering Survey (APPLES) (TR-10-01)," Center for the Advancement for Engineering Education., Seattle, WA, 2010.

[21] H. J. Passow, "Which ABET Competencies Do Engineering Graduates Find Most Important in Their Work?," *Journal of Engineering Education,* vol. 101, no. 1, pp. 95-108, 2012.

[22] B. G. Tabachnick and L. S. Fidell, Using multivariate statistics (5th ed.), Allyn & Bacon/Pearson Education, 2007.

[23] M. D. Evans and M. E. Jordan, "How writing for the public provides affordances and constraints in enacting expert identity for undergraduate engineering students," in *ASEE Annual Exposition & Conference*, Tampa, FL., 2019.

[24] B. Ponomariov, E. Welch and J. Melkers, "Assessing the outcomes of student involvement in research: educational outcomes in an engineering research center," *Research Evaluation,* vol. vol, no. 4, pp. 313-322, 2009.

[25] M. M. Pop, P. Dixon and C. Grove, "Research Experiences for Teachers (RET): Motivation, Expectations, and Changes to Teaching Practices due to Professional Program Involvement," *Journal of Science Teacher Education,* vol. 21, no. 2, pp. 127-147, 2010.

[26] C. Faber, E. Hardin, S. Klein-Gardner and L. Benson, "Development of teachers as scientists in research experiences for teachers programs," *Journal of Science Teacher Education,* vol. 25, no. 7, pp. 785-806, 2014.

[27] B. L. Ponomariov and P. C. Boardman, "Influencing scientists' collaboration and productivity patterns through new institutions University research centers and scientific and technical human capital,," *Research Policy,* vol. 39, no. 5, pp. 613-624, 2010.

[28] M. R. Blanchard, S. A. Southerland and E. M. Grang, "No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers," *Science Education,* vol. 93, no. 2, pp. 322-360, 2008.

[29] R. J. Miranda and J. B. Damico, "Science teachers' beliefs about the influence of their summer research experiences on their pedagogical practices," *Journal of Science Teacher Education,* vol. 24, no. 8, pp. 1241-1261, 2013.

[30] A. L. Zydney, J. S. Bennett, A. Shahid and K. W. Bauer, "Impact of Undergraduate Research Experience in Engineering," *Journal of Engineering Education,* vol. 91, no. 2, pp. 151-157, 2002.

[31] W. Aung, "Undergraduate and graduate education activities of current engineering research centers.," ERC Education Assessment and Dissemination Task Group, 2006.

[32] S. H. Russel, "Evaluation of NSF support for undergraduate research opportunities: Survey of NSF program participants," National Science Foundation, Directorate for Engineering, Division of Engineering Education and Centers. SRI International, Menlo Park, CA, 2005.

[33] S. H. Russell and M. P. Hancock, "Evaluation of the Research Experience for Teachers (RET) program: 2001-2006," National Science Foundation, Directorate for Engineering, Division of Engineering Education and Centers. SRI International, Menlo Park, CA., 2006.

Appendix A – Pairwise Pearson correlations between items associated with understanding the ERC

		present				prior				
		fields	real- world	engineering problem	career	mission	fields	real- world	engineering problem	career
present	mission	.784	.650	.661	.693	.361	.403	.389	.372	.392
	fields		.665	.713	.680	.389	.444	.337	.322	.301
	real-world			.667	.584	.407	.262	.427	.279	.238
	engineering problem				.696	.376	.368	.399	.413	.311
	career					.329	.323	.375	.381	.486
prior	mission						.587	.576	.536	.468
	fields							.664	.756	.686
	real-world								.743	.667
	engineering problem									.742

All correlations are significant .01 level.

Codes:
mission: I understand the mission of the ERC
fields: I understand the primary field(s) involved in the ERC
real-world: I understand the connection between the ERC field(s) of study and how the ERC helps people address real world issues
engineering problem: I understand the engineering problems associated with the ERC field(s) of study
career: I understand the career pathway(s) associated with the ERC field(s) of study

Appendix B – Pairwise Pearson correlations between items associated with skillset

	WR	VI	NI	NP	DP	SA	IF	CO	RQ	AD	SP	CD	KR	EQ	MC	NX	WI	ET	LS	RL	CL	MT
OP	.660	.627	.361	.471	.516	.711	.528	.322	.461	.551	.530	.478	.397	.459	.406	.388	.321	.344	.326	.434	.435	.426
WR		.511	.394	.431	.469	.546	.460	.346	.493	.404	.418	.299	.289	.270	.334	.298	.271	.251	.293	.430	.438	.356
VI			.294	.485	.457	.553	.498	.398	.429	.418	.431	.393	.386	.317	.384	.367	.272	.262	.281	.190	.274	.264
NI				.573	.344	.470	.426	.524	.352	.315	.426	.397	.304	.245	.401	.294	.322	.306	.304	.339	.347	.375
NP					.532	.568	.642	.581	.471	.420	.553	.475	.473	.329	.401	.459	.385	.434	.324	.468	.507	.434
DP						.548	.656	.461	.459	.389	.500	.445	.388	.405	.355	.459	.369	.410	.403	.430	.459	.386
SA							.672	.489	.575	.587	.660	.460	.434	.443	.503	.440	.519	.416	.465	.491	.542	.586
IF								.708	.543	.563	.710	.628	.589	.478	.442	.523	.587	.526	.437	.517	.552	.584
CO									.409	.477	.625	.604	.556	.426	.481	.441	.571	.537	.422	.460	.514	.424
RQ										.686	.637	.548	.570	.402	.574	.631	.500	.548	.439	.536	.585	.435
AD											.713	.804	.681	.587	.626	.658	.588	.501	.446	.552	.558	.594
SP												.723	.716	.548	.568	.669	.642	.632	.501	.633	.713	.609
CD													.790	.643	.618	.628	.585	.626	.454	.547	.577	.599
KR														.669	.591	.677	.639	.591	.460	.567	.627	.489
EQ															.438	.431	.519	.500	.318	.485	.529	.439
MC																.668	.518	.529	.468	.493	.556	.511
NX																	.583	.496	.526	.601	.654	.478
WI																		.638	.617	.630	.601	.578

ET																		.643	.630	.618	.557
LS																			.565	.530	.533
RL																				.885	.653
CL																					.601

All correlations are significant .01 level.

Codes:
OP: Oral presentation
WR: Writing technical reports, conference proceedings, or journal publications
VI: Creating visual displays such as posters or prototypes
NI: Networking with industry
NP: Networking with my peers
DP: Defending a position as part of a discussion
SA: Speaking appropriately to different audiences
IF: Incorporating feedback
CO: Collaborating with others
RQ: Formulating research questions
AD: Analyzing data (e.g., experiment results)
SP: Solving problems
CD: Collecting data from lab experiments, field observations, or web searching
KR: Keeping a record of research activities
EQ: Using lab equipment
MC: Making connections between classroom learning and research
NX: Determining the next step in a research project
WI: Working independently
ET: Conducting research in an ethical and responsible manner
LS: Providing leadership on projects
RL: Finding relevant literature
CL: Making connections between existing literature and research
MT: Managing time

Appendix C – Pairwise Pearson correlations between items associated with mentorship experience

	direct	training	feedback	encourage	inspire	share	demonstrate	goal	career	model	help	profession	challenge
advice	.642	.686	.674	.754	.788	.740	.637	.564	.584	.538	.639	.660	.440
direct		.791	.764	.669	.722	.645	.527	.622	.409	.411	.594	.505	.507
training			.797	.660	.739	.599	.473	.539	.415	.403	.503	.517	.629
feedback				.808	.748	.647	.477	.557	.564	.463	.660	.606	.576
encourage					.811	.712	.501	.642	.661	.559	.731	.701	.524
inspire						.595	.453	.583	.583	.496	.635	.616	.501
share							.813	.643	.614	.569	.757	.820	.611
demonstrate								.545	.467	.524	.702	.762	.401
goal									.533	.573	.779	.681	.473
career										.613	.579	.539	.577
model											.599	.626	.557
help												.845	.456
profession													.490

All correlations are significant .01 level.

Codes:
advice: advice that supports your future plans
direct: direction on your research project
train: training to support independent work
feedback: feedback that is constructive
encourage: encouragement to strive for success
inspire: inspiration to pursue a career in a STEM-related field:
share: shared knowledge and expertise
demonstrate: demonstrated knowledge and expertise
goal: established my project goals:
career: discussed my career goals:
model: provided me with a role model:
help: helped me when needed:
profession: displayed professionalism:
challenge: challenged me to extend my ability

Appendix D – Pairwise Pearson correlations between items associated with climate of inclusion (Note: only two decimal places shown due to space considerations.)

	PI	PE	ID	II	IE	VD	VI	VE	FA	CL	EQ	EN	DI	BI	EX	FM	ME	ST	WR	NW	IP
PD	.82	.82	.60	.71	.64	.88	.75	.67	.45	.78	.73	.42	.02	-.15	-.17	.46	.29	.28	.26	.26	.37
PI		.89	.59	.70	.67	.78	.87	.78	.51	.82	.85	.41	-.05	-.16	-.27	.49	.36	.37	.37	.40	.37
PE			.63	.68	.66	.81	.86	.84	.58	.89	.84	.48	-.01	-.20	-.20	.50	.37	.39	.32	. 30	. 38
ID				.78	.73	.67	.63	.56	.39	.53	.49	.50	.03	-.14	-.12	.32	.15	.23	.21	.21	.36
II					.84	.77	.75	.70	.39	.62	.59	.51	-.09	-.09	-.17	.37	.23	.28	.24	.29	.34
IE						.68	.65	.77	.33	.60	.53	.55	-.05	-.12	-.09	.30	.14	.26	.20	.31	.34
VD							.75	.68	.43	.73	.72	.45	.02	-.18	-.15	.39	.25	.24	.24	.28	.42
VI								.80	.54	.78	.81	.43	-.14	-.20	-.31	.46	.35	.37	.40	.44	.44
VE									.50	.76	.73	.46	-.17	-.26	-.20	.40	.25	.24	.23	.31	.17
FA										.65	.61	.09	-.20	-.36	-.46	.46	.34	.27	.27	.23	.22
CL											.85	.38	-.09	-.28	-.31	.52	.29	.31	.29	.31	.11
EQ												.40	-.02	-.20	-.32	.41	.24	.29	.33	.30	.21
EN													.15	.08	.11	.21	.15	.11	.15	.20	.36
DI														.91	.96	-.23	-.33	-.20	-.35	-.21	-.10
BI															.89	-.33	-.24	-.30	-.32	-.21	-.24
EX																-.32	-.35	-.33	-.41	-.30	-.31
FM																	.73	.62	.57	.56	.72
ME																		.68	.69	.66	.67

ST																			.70	.68	.73
WR																				.86	.66
NW																					.69

Dark grey shading: Pearson correlation is significant at 0.05 level
Light grey shading: Pearson correlation is significant at 0.01 level

Codes:
PD: implemented policies and practices that reinforce a commitment to diversity
PI: implemented policies and practices that reinforce a commitment to inclusion
PE: implemented policies and practices that reinforce a commitment to equity
ID: stressed the importance of learning about diversity
II: stressed the importance of learning about inclusion
IE: stressed the importance of learning about equity
VD: made diversity an essential value of the ERC
VI: made inclusion an essential value of the ERC
VE: made equity an essential value of the ERC
FA: treated all students fairly, irrespective of their membership in diverse groups (i.e., age, race/ethnicity, religion, gender, etc.)
CL: created a positive climate for those who belong to diverse groups
EQ: provided equal opportunities and access for all to participate in ERC activities
EN: been encouraged by the ERC to participate in diversity and inclusion events
DI: heard disparaging remarks made about a member or members of the ERC due to their membership in a diverse group(s)
BI: witnessed or have been a victim of bias within the ERC
EX: witnessed or have been a victim of exclusion within the ERC
FM: I have felt accepted while interacting with ERC faculty member(s)
ME: I have felt accepted while interacting with ERC mentors
ST: I have felt accepted while interacting with ERC staff members
WR: I have felt accepted while interacting with ERC peers within a work-related situation
NW: I have felt accepted while interacting with ERC peers within a non-work-related situation
IP: I have felt accepted while interacting with ERC industry partners