ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #37338

Developing common qualitative tools for cross ERC education program evaluation

Zhen Zhao

Zhen Zhao is a Ph.D. student in The Polytechnic School at Arizona State University. His research interests include engineering student mentorship and leadership development, engineering research center education and diversity impact evaluation, and engine

Megan O'donnell (Research Professional)

Marcus Lyra

Engineers are motivated by innovation and new ideas, many scholars have spent their lives in finding and suggesting effective ways of supporting long-life learning in engineering (from K-12 to professional engineers). I am also one of them, specifically, I am passionate about how engineering educators become educators, how they seek teaching-related professional development, and how it is translated into effective engineering courses, using quantitative, qualitative, and mixed-methods analysis.

Michelle Jordan (Assistant Professor)

Wilhelmina C. Savenye (Professor Emeritus) (Arizona State University)

Gillian Roehrig (Professor)

Developing common qualitative tools for cross ERC education program evaluation

Abstract

National Science Foundation (NSF) funded Engineering Research Centers (ERC) are required to develop and implement education and outreach opportunities related to their core technical research topics to broaden participation in engineering and create partnerships between industry and academia. Additionally, ERCs must include an independent evaluation of their education and outreach programming to assess their performance and impacts. To date, each ERC's evaluation team designs its instruments/tools and protocols for evaluation, resulting in idiosyncratic and redundant efforts. Nonetheless, there is much overlap among the evaluation topics, concepts, and practices, suggesting that the ERC evaluation and assessment community might benefit from having a common set of instruments and protocols. ERCs' efforts could then be better spent developing more specific, sophisticated, and time-intensive evaluation tools to deepen and enrich the overall ERC evaluation efforts. The implementation of such a suite of instruments would further allow each ERC to compare its efforts to those across other ERCs as one data point for assessing its effectiveness and informing its improvement efforts. Members of a multi-ERC collaborative team, funded by the NSF, have been leading a project developing a suite of common instruments and protocols which contains both quantitative and qualitative tools. This paper reports on the development of a set of qualitative instruments that, to date, includes the following: (a) a set of interview/focus group protocols intended for various groups of ERC personnel, centered around five common topics/areas, and (b) rubrics for summer program participants' verbal poster/presentations and their written poster/slide deck presentation artifacts. The development process is described sequentially, beginning with a review of relevant literature and existing instruments, followed by the creation of an initial set of interview questions and rubric criteria. The initial versions of the tools were then pilot-tested with multiple ERCs. Feedback sessions with education/evaluation leaders of those piloting ERCs were then conducted, through which further revision efforts were made.

Key Words

Evidence-based practice paper, ERC Evaluation, Qualitative Tools,

Introduction

75 Engineering Research Centers (ERCs) across the nation have been funded by the National Science Foundation (NSF) since 1985; 14 are currently operating [1]. These ERCs have played a big part in not only advancing engineering and technology but also integrating research, education, and workforce development [1]. NSF requires each ERC to provide educational and professional development opportunities for professionals, post-doctoral researchers, graduate students, undergraduate students, K-14 teachers, and K-12 students. At the same time, data-driven approaches are recommended to evaluate and track the performance and impacts of these opportunities [2]; findings are required as part of the center's annual report and site visit presentations. The responsibility falls on each ERC to develop and implement an evaluation

plan. The effort, in general, is led by center education directors/leadership collaborating with external evaluators.

Each center is given the autonomy to select preferred evaluation techniques and tools (quantitative and/or qualitative). This flexibility, to date, has resulted in ERCs' isolated evaluation efforts, especially in developing evaluation tools, which causes redundant resource spending. The educational and professional development opportunities provided by each ERC share similar structures, settings, missions, personnel, and logistics [3]. Such siloed efforts also make cross-ERC comparisons essentially unfeasible. Limited work has responded to these limitations despite NSF's encouragement for cross-center collaboration in program evaluation. A collaborative consortium initiated direct, conscientious work to combat these ongoing issues during ERC evaluation and was awarded a 3-year NSF grant to conduct these activities [4]. This paper reports the year-one achievement in qualitative tools development.

Project Description

The NSF-funded project aims to provide a new approach to assessing the effectiveness of education and professional development programs within and across ERCs, by developing a suite of common evaluation tools that can be used by all ERCs and similarly structured large-scale science and technology research centers [4]. The development of the suite of common evaluation tools will not only reduce ERC external evaluators' workload in creating tools for individual centers, and enhance the evaluation quality and efficiency, but also make it possible to expand and share evaluation findings across ERCs. These efforts will enhance collaboration across ERCs and provide each center with insights on how to improve its educational and professional development programs.

The Multi-ERC Instrument Inventory (MERCII) is an expanding suite of common evaluation tools intended to be a freely shared resource used by the ERC community and throughout the greater STEM research community. The MERCII includes a streamlined quantitative instrument, a set of complementary qualitative protocols, an updated evaluation section for NSF ERC Best Practices Manual [2], a supplemental evaluator toolbox, and an independent mercii.org online platform (to help ERCs disseminate the materials). The different sets of evaluation tools are complementary to each other and combined will cover all the measurements suggested in the Evaluation section of the Best Practices Manual. This paper focuses on qualitative tools.

The collaborative consortium that led the project consists of education directors, diversity leaders, and external evaluators from six different ERCs at varying stages of existence (one at year two, two at year five, two at year seven, and one just finished year ten), plus experienced quantitative and qualitative researchers in engineering education. The consortium is set to collectively utilize the expertise, experiences, and resources of its members and the partner ERCs. Some prior cross-ERC activities designed and implemented by the consortium include outreach activities, invited industry talks, shared summer programs, and conference/meeting workshops [5-6].

Per the recommendation of the Best Practices Manual, evaluators have been involved from the beginning of this project and have been key in helping ensure that ERC evaluation protocols align with center goals. Iterative feedback has been sought from the wider ERC community throughout the development process. Formative and impact assessment protocols are provided to help inform feedback loops with the centers' education teams and ultimately enhance the impact of the center on all participants.

Qualitative Tools

The project began with the effort to develop a streamlined quantitative instrument that can be used across ERC participant groups (e.g., faculty members, year-long research assistants, and summer interns) [4, 7-8]. A quantitative instrument alone will not fit all the evaluation needs for every ERC as the quantitative instrument does not provide detailed information behind the scale scores. This challenge suggests a need to complement the quantitative instrument with qualitative tools that are effective for collecting descriptive and explanatory data.

The complementary qualitative tools being developed include protocols and rubrics for interviews and focus groups, poster assessments, and presentation observations. The development of the qualitative protocols followed an iterative waterfall process which was also utilized while constructing the quantitative instrument [7, 8]. Steps included synthesizing existing protocols, referencing supplemental literature, team discussion, and multiple iterations of testing and revising. The validity of the evaluation tools has been and will be achieved by 1) thoroughly comparing our creations with already in-use qualitative ERC evaluation tools, 2) repetitively having experts review our work, 3) iteratively testing updated versions across multiple centers, settings, and population groups. 4) constantly incorporating feedback from evaluators and participants. This paper introduces the development of two qualitative tools that have achieved the furthest progress: interview/focus-group protocols and poster /presentation rubrics.

These qualitative protocols are designed to provide increased flexibility, allowing for greater variation in implementation compared to the quantitative instrument. Unlike the quantitative instrument, qualitative tools are usually tailored and implemented by evaluators to accommodate the needs of each center. The project's goal of developing qualitative tools is to provide a template to allow for some level of consistency across ERCs, in addition to allowing the potential for cross-center comparisons, should ERC education leaders wish to conduct such analyses.

Interview/Focus-group protocols

The first qualitative assessment tool developed was interview/focus-group protocols. The consortium started the process in spring 2020 by collecting all the interview and focus-group protocols used by three different partner ERCs. All questions were extracted and compared across the protocols. Five common categories emerged after aligning the questions. These five common categories are listed in table 1.

In fall 2020, the consortium conducted extensive reviews and commentaries, as well as frequent discussions, to determine a set of initial questions that would be included in the protocol. All prior extracted questions were listed per category in a spreadsheet. All consortium members were asked to rate each question with "essential (E)" or "supplemental (S)" and then to rank the priority of each question from "1" to "5" with the lower number being more prioritized. Since the consortium team would be discussing their ratings and rationales extensively, ratings were not done anonymously. The questions within a category were then sorted from the most essential counts to the least essential counts with the rank as the tiebreaker if the questions had the same essential counts. All consortium members met together to discuss the ratings for each question. Consortium members were allowed to change their question ratings as the discussion proceeded. In the end, questions marked as essential by all team members were determined to be kept in the protocols, and the ones with no scores of essential were eliminated. A few questions that fell in between were determined to be kept temporarily and to need further investigation. Meanwhile, equivalent or duplicated questions were combined and reworded.

Table 1. *Five common categories in interview/focus-group protocols*

Category
Impact on Skills
Culture of Inclusion
Mentoring Experience
STEM-related Future Plans
Program Satisfaction

A sub-team of the consortium continued to update the protocols throughout spring 2021 by searching relevant ERC qualitative evaluation literature and resources to seek applicable questions that could complement the initial set of the questions. This sub-team also worked on rewording questions, further examining the temporarily kept questions, and drafting a protocol introduction, instructions, and transitions between categories. The sub-team reported work progress to the whole consortium team and asked for feedback periodically. All questions were also reviewed and reworded by the team to assure compatibility between the interview and focus group protocols. Instructions and notes for applying the protocols in both settings were also included.

The interview/focus-group protocols developed to date include a set of three protocols.: The baseline protocol, the Research Experiences for Teachers (RET) protocol, and the mentor protocol. The baseline interview/ focus-group protocol includes questions that apply to the widest range of participant groups possible (e.g., post-doctoral researchers, graduate students, undergraduate students, and summer program participants (Research Experiences for Undergraduates (REU), Young Scholars Program (YSP), Research Experience and Mentoring (REM), and Research Experiences for Veteran Undergraduates (REV)). The RET interview/focus-group protocol contains all the questions from the basic protocol (except the

questions under the "STEM-related future plans" category) plus questions regarding lesson plan development. The mentor interview/focus-group protocol involves questions only regarding mentors' mentoring experiences due to the extensiveness of the mentoring experience-related questions.

Table 2. *Example questions in the current version of interview/focus-group protocols*

Category	**Example Questions**
Impact on Skills	• *Tell me about some of the things you've learned through your participation in [RC*] [REU/RET/YSP/REM/REV Program].* • *Please provide examples of how you will use these things you've learned moving forward.*
Culture of Inclusion	• *In your experience, did the [RC] create a culture of inclusion? Please give some examples.*
Mentoring Experience	• *Please provide some examples of the mentoring you received from your mentor(s) during your overall [RC] experience.*
STEM-related Future Plans	• *Has your overall [RC] experience impacted your future plans? If so, please provide examples.*
Program Satisfaction	• *If you found your [RC] experience beneficial, please describe the most beneficial aspect.*
Research project Meetings	• *Most [RC] faculty members host weekly or bi-weekly lab meetings across research projects. Do you participate in these meetings? [Follow-up] If so, how useful are lab meetings for your learning and your research progression?*
RET Lesson Plan	• *Please describe anything that was helpful and/or challenging in developing your lesson. Do you have any recommendations for improvement?*

*RC (i.e., Research Center) is used in the instruments in order to facilitate the use of the protocols beyond ERCs.

The first version of the baseline protocols was piloted in both focus-group and interview settings with two ERCs and two NSF-funded non-ERC STEM research centers during the summer of 2021. Another two partner ERCs used their own versions that shared great similarities with the protocols. The RET protocol and mentor protocol were not piloted in this round because of timing and budget restrictions. The revision of the baseline protocol started at the beginning of fall 2021 with input from consortium members who piloted the protocol, including the two centers that used their own versions. The biggest change was cutting the number of questions from 16 to 13 and clarifying how to select questions for a shorter interview/focus group. Other updates included reducing the word counts in the questions, incorporating two questions from the partner ERCs' protocols into the baseline questions, and

editing the protocol implementation instructions. The revising activities occurred throughout the academic year. The most recent change made was the addition of questions asking about the impact and experiences of ERC project research meetings. These questions were added in order to address NSF's critique at a partner ERC's site visit. The revised protocol is ready to test again starting from 2022 summer. Example questions under the five common categories, RET lesson plan, and ERC project research meeting in the current version of protocols are listed in table 2.

Poster/Presentation Rubrics

The rubrics are designed to be used in assessing poster/presentation slide artifacts as well as verbal presentations for individuals participating in REU, RET, and YSP programs. The initial version was created in spring 2021 and piloted in summer 2021. Initial elements of the rubric and their descriptions were based on the collective experience and comparison of past tools used by team members with close ties to the six partner ERCs. The rubrics were adopted by four different ERCs. Evaluators, education teams, and summer program mentors were invited to use the rubrics to assess summer program participants in their research centers. They then were invited to consortium meetings to provide feedback on their experiences and suggestions at the beginning of 2021 Fall. Further revisions followed. Additionally, the consortium members applied the updated rubrics to first independently assess posters and verbal presentations archived from the 2021 summer data collection, before comparing and negotiating ratings and notes, making additional modifications to further enhance the tools by elaborating and clarifying the elements. This process was time-intensive, as the independent assessments and discussions/negotiations were conducted for six presentations and posters from multiple participant groups (YSP, REU, and RET) across two centers, including the Center for Bio-mediated and Bio-inspired Geotechnics (CBBG), and the Quantum Energy and Sustainable Solar Technologies (QESST) programs. Scoring bias was further accounted for by the qualitative team being comprised of faculty members, educators, evaluators, and graduate students from different fields and with diverse gender, racial/ethnic, and educational backgrounds.

The modified version of these rubrics includes four columns (yes, somewhat, no, and notes/suggestions) that assess the degree to which each component was delivered by participants. Suggestions are encouraged in those instances in which a "somewhat" or "no" was marked. The rubrics were designed to assess the impact as well as provide formative feedback to program participants before they present at site visits or other meetings in future dissemination efforts. While these rubrics were designed primarily for the evaluators to complete, they may also be completed by members of the Research Center education team, and/or mentors (faculty or graduate students). The top of the rubric asks the observer to indicate their role in the Research Center and also identify the role of the person being observed (e.g., REU, RET, YSP).

The rubrics include a core set of elements for all program participants as well as specific elements only assessed in those participating in the RET program. The common elements assessed in all groups include two sections, one for the visual aspects of the poster/presentation slides, and one for the verbal aspects of the actual presentation.

Visual Presentation Rubric

Figure 1. A copy of Visual Presentation Rubric – Visual & Written Elements

Visual Presentation Rubric - Visual & Written Elements

Date of assessment _____ Presentation format: poster slides other_____

Observer _____ Role in the RC_____

Presenter _____ Role in the RC _____

Context of the assessment: end of summer presentation lab meeting other_____

Visual & Written Elements				
	Yes	Some what*	No*	Notes/Comments
Readability				
• easy to read				
• carefully edited				
• font size & color readable				
• uses space appropriately (e.g., text not too dense, effective use of bullets, well organized, clear headings)				
• ADA compliant (e.g., color blind accessibility, text captions, audio, screencast for recording)				
Written Content				
• appropriately concise (necessary information only)				
• relevant information only				
• understandable by a public audience				
Visual Representations (e.g., photos, graphs, tables, sketches)				
• canonical usage as defined by engineering and/or science communities (e.g., figures labeled, units included, axes identified)				
• enhance the text; added for a specific purpose				
• graphics are high quality				

* If you mark "somewhat" or "no", please leave notes.

Notes:

This material is based upon work primarily supported by the National Science Foundation under award EEC 2023275. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

Visual and Written Elements. The visual and written elements include an assessment of the degree to which the poster/slides are 1) easy to read (e.g., font size and color readable, spaced appropriately, and ADA compliant), 2) content is appropriately concise, relevant, and edited appropriately, and 3) visual representations are effective (labeled photos, graphs, charts, and tables as considered acceptable by engineering and/or science communities, increases understandability of content, and is easy to interpret). A copy of the Visual Presentation Rubric - Visual and Written Elements is provided in figure 2.

Research Elements. The research elements include an assessment of the extent to which the poster/presentation slides demonstrate the following components: background (e.g., connects the research project to literature), research purpose (e.g., defines the problem and/or research questions), method (e.g., describes lab processes, procedures, experiments, data collection, and/or analysis), results, conclusions (e.g., connects to literature, identifies next steps, future research), and citations. Brief descriptions of each component are provided in the rubric to help guide those who are performing the assessment.

RET Lesson Elements. This component of the rubric includes assessment of the following: Lesson development and description, and lesson implementation, both of which would be ideally completed by someone from the education team within the Research Center. The lesson development section includes the following: appropriate learning goals/outcomes, alignment between lesson activities and assessment strategies with learning goals, and explicit attention to culturally and linguistically diverse students, as well as those with special needs/exceptionalities. The lesson implementation sub-section includes the extent to which the lesson is adapted for remote delivery (when appropriate) and the action plan identified in the implementation.

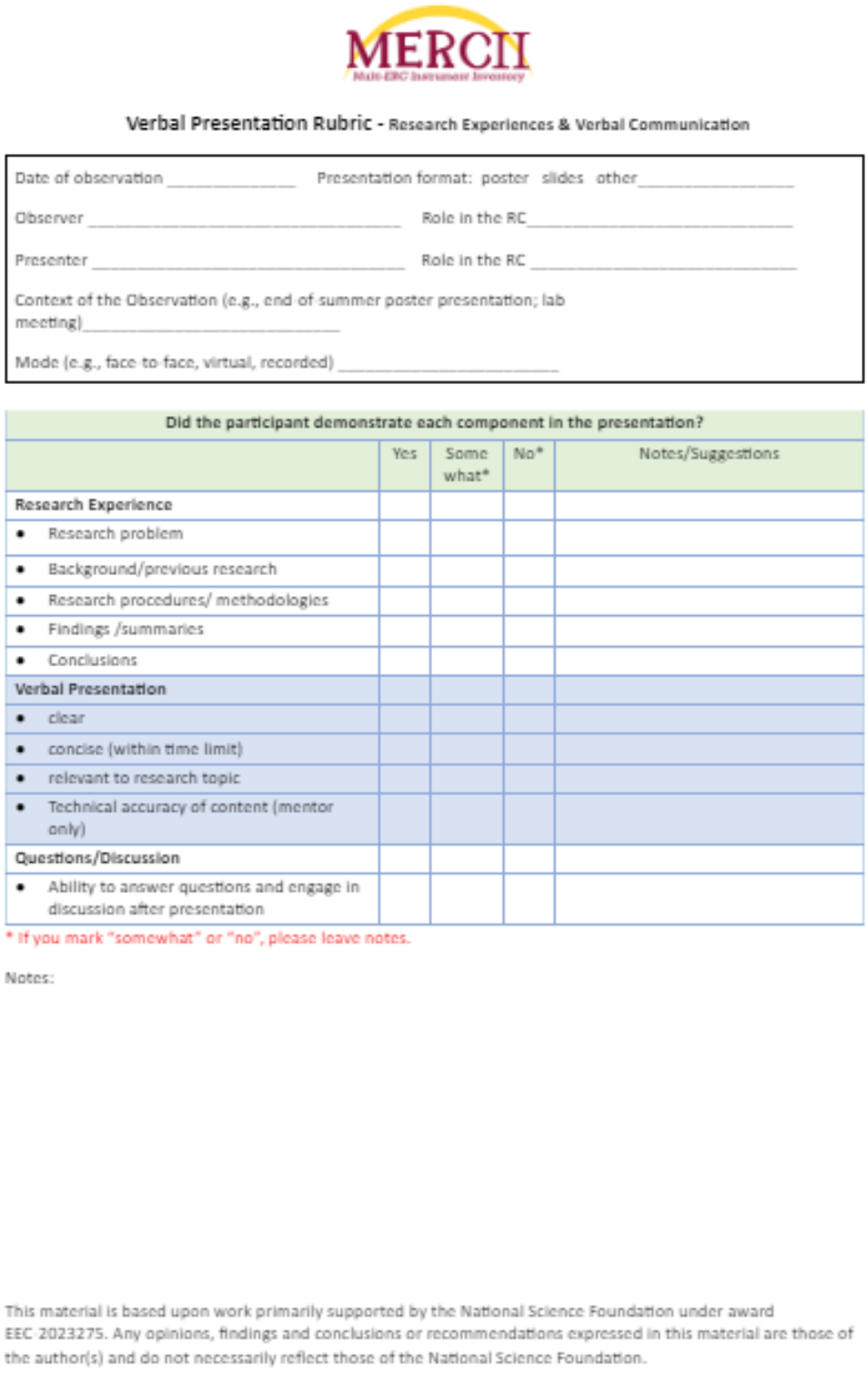
Verbal Presentation Rubric

The verbal presentation rubric is a separate section of the rubric used to assess the verbal presentation and is divided into the following three subsections: Research Experiences, Verbal Communication, and RET Lesson Elements. A copy of the Verbal Presentation Rubric – Research Elements & Verbal Communication is provided in figure 2.

Research Experiences. This subsection of the rubric includes an assessment of whether participants demonstrated the following research components: research problem, background/previous research, procedures/methodologies, and conclusions.

Verbal Communication. This subsection includes an assessment of the clarity and conciseness and relevance of the presentation to the research topic and an assessment of the presenter(s)' ability to answer/field questions at the end of the presentation as well as engage in discussion with those who are fielding questions/comments at the end. There is a section called "technical accuracy of the content," which is only to be assessed by the participants' mentor(s).

Figure 2. A copy of Verbal Presentation Rubric – Research Experiences & Verbal Communication



MERCII
Multi-ERC Instrument Inventory

Verbal Presentation Rubric - Research Experiences & Verbal Communication

Date of observation _____ Presentation format: poster slides other_____

Observer _____ Role in the RC_____

Presenter _____ Role in the RC _____

Context of the Observation (e.g., end of summer poster presentation; lab meeting)_____

Mode (e.g., face to face, virtual, recorded) _____

Did the participant demonstrate each component in the presentation?				
	Yes	Some what*	No*	Notes/Suggestions
Research Experience				
• Research problem				
• Background/previous research				
• Research procedures/ methodologies				
• Findings /summaries				
• Conclusions				
Verbal Presentation				
• clear				
• concise (within time limit)				
• relevant to research topic				
• Technical accuracy of content (mentor only)				
Questions/Discussion				
• Ability to answer questions and engage in discussion after presentation				

* If you mark "somewhat" or "no", please leave notes.

Notes:

This material is based upon work primarily supported by the National Science Foundation under award EEC 2023275. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

RET Lesson Elements. The RETs will be assessed on their ability to present their lesson plan development. Specifically, they will be assessed on whether they clearly define/describe the following elements: ABET/NGSS/State/Tribal or other standards, student learning outcomes/objectives, implementation plan, assessment strategies, mindfulness of culturally and linguistically diverse students, as well as those students with special needs/exceptionalities, adaptation to deliver the lesson remotely (if relevant), and connectivity of lesson with teachers' real-world lab experience and student learning experiences.

Next Steps

Two tasks have been planned for the next steps: 1) finishing the creation of the first version of the RET lesson plan rubrics and 2) developing an ERC virtual experience assessment tool (to be used with participants who are participating remotely, for instance, during Covid restrictions). The RET lesson plan rubrics contain two independent rubrics: a lesson plan evaluation rubric and a lesson plan implementation observation rubric. The former assesses the quality of the instructional plan that the RET teachers develop, and the latter evaluates the quality of teachers implementing the lesson plan in the classroom. Both rubrics have been initiated and are currently under development. ERC virtual experiences will be the next topic covered. The type of tools is yet undecided.

Acknowledgments

This work is supported by the National Science Foundation Grant EEC-2023275. Any opinions, findings, 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. We would also like to thank the leadership team, education team, and the evaluation team of all partner ERCs, for their support and participation in knowledge sharing, data collection, and offering constructive feedback.

References

[1] National Science Foundation, "Engineering Research Centers," National Science Foundation, [Online]. Available: https://nsf.gov/eng/eec/erc.jsp. [Accessed 9 February 2022].

[2] Court, "Best Practices Manual/Chapter4: Education Programs/Section 4.6: Assessment and Evaluation," National Science Foundation, 23 10 2014. [Online]. Available: https://erc-assoc.org/best_practices/section-46-assessment-and-evaluation. [Accessed 9 Feburary 2021].

[3] National Science Foundation, "Education Core Mission," National Science Foundation, [Online]. Available: https://erc-assoc.org/content/education-core-mission. [Accessed 9 February 2022].

[4] Z. Zhao, A. Carberry, M. Jordan, J. Larson, W. Savenye, W. Barnard, M. O'Donnell, C. Argenti, C. Barr, A. Godwin and G. Roehrig, "Design and Development: NSF Engineering Research Centers Unite: Developing and Testing a Suite of Instruments to Enhance Overall Education Program Evaluation," in *American Society of Engineering Education (ASEE) Annual Conference & Exposition*, Long Beach, CA, July 2021.

[5] J. S. Larson, M. O'Donnell, W. M. Barnard, A. Cook-Davis, A. R. Carberry and S. B. Adams, "Three ERCs and a National Network Node: Assessing Engineering Outcomes for Middle School Students Across a Joint Outreach Event," in *American Society of Engineering Education (ASEE) Annual Conference & Exposition*, Tempa, FL, June 2019.

[6] J. Larson, A. Carberry, M. Jordan, A. Cook-Davis and M. O'Donnell, Writers, *The ASU Tri-ERC Consortium: Establishing a common set of tools for evaluating educational programs within and across ERCs.* [Performance]. Workshop at the NSF ERC Biennial Meeting, 2019.

[7] Z. Zhao, A. Carberry, W. Barnard, A. Cook-Davis, M. Jordan, J. Larson, M. O'Donnell, and W. Savenye, "Creating Common Instruments to Evaluate Education and Diversity Impacts across Three Engineering Research Centers," in *Frontiers in Education (FIE) Annual Conference*, Cincinnati, OH, October 2019.

[8] Z. Zhao, A. Carberry, A. Cook-Davis, J. Larson, M. Jordan, W. Barnard, M. O'Donnell and W. C. Savenye, "Streamlining the process of evaluating the education and diversity impacts across Engineering Research Centers," in *American Society of Engineering Education (ASEE) Annual Conference & Exposition*, Vancouver, B.C., 2020.