

Evaluation of the Second Edition of *Entering Research*: A Customizable Curriculum for Apprentice-Style Undergraduate and Graduate Research Training Programs and Courses

Janet L. Branchaw,^{***} Amanda R. Butz,[†] and Amber R. Smith[†]

[†]Wisconsin Institute for Science Education and Community Engagement and [†]Department of Kinesiology, University of Wisconsin–Madison, Madison, WI 53706

ABSTRACT

The second edition of *Entering Research* (ER) is a collection of customizable active-learning activities, resources, and assessment and evaluation tools for use in undergraduate and graduate research training programs and courses. Results from two design and development research studies examining the effectiveness of the second edition of the ER curriculum and a 2-day ER facilitator training workshop are reported. Pilot testing of the second edition of the curriculum at 20 sites across the country (42 unique implementations) with 78 facilitators and 565 undergraduate and graduate research trainees provides evidence that the ER activities are clear and complete and that they were effective in helping trainees gain knowledge or improve their ability to do research. Overall, research training program directors and trainees were satisfied with courses and workshops that incorporated activities from ER. Likewise, evaluation data from four ER facilitator training workshops showed that participants valued the workshop and reported significant gains in confidence in their ability to successfully develop and implement a custom ER curriculum. Together, these results provide evidence that the ER curriculum and training workshop warrant further efficacy, effectiveness, and scale-up research.

INTRODUCTION

The traditional, apprentice-style research learning experience, in which students work one-on-one with a research mentor, is the primary model used to train future biology researchers and future researchers in science, technology, engineering, mathematics, and medicine (STEMM) generally. Though course-based research experiences (CUREs) are becoming more common and are broadening access to research learning experiences (National Academies of Sciences, Engineering, and Medicine [NASEM], 2015), students who choose to pursue a career in research will participate in apprentice-style research as part of their training. Therefore, providing high-quality, accessible apprentice-style research experiences is critical to our efforts to develop and diversify the next generation of STEMM researchers.

Apprentice-style research experiences have been shown to contribute to the persistence of undergraduate and graduate students in biology and STEMM generally, especially for students who are members of underrepresented groups (Lopatto, 2007; Chemers *et al.*, 2011; Eagan *et al.*, 2013; Chang *et al.*, 2014; Linn *et al.*, 2015). They are high-impact learning experiences that can be transformative, altering a student's career trajectory toward or away from research, and are highly valued by STEMM graduate program admissions committees (Sabitini, 1997; Hathaway *et al.*, 2002; Seymour *et al.*, 2004; Hunter *et al.*, 2007; Carter *et al.*, 2009; Junge *et al.*, 2010; Laursen *et al.*, 2010; Craney *et al.*, 2011; Thiry *et al.*, 2011). Consequently,

Gil Marbach-Ad, *Monitoring Editor*

Submitted Apr 9, 2019; Revised Jan 14, 2020;
Accepted Jan 22, 2020

CBE Life Sci Educ March 1, 2020 19:ar11

DOI:10.1187/cbe.19-04-0073

*Address correspondence to: Janet Branchaw
(branchaw@wisc.edu).

© 2020 J. L. Branchaw *et al.* CBE—Life Sciences Education © 2020 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

apprentice-style research experiences often serve as gateways to STEMM graduate training programs and research careers.

Structured research training programs (e.g., graduate training programs, research experiences for undergraduates programs, research methods courses) provide just-in-time support for cohorts of trainees engaged in apprentice-style research and help to disperse the power differential inherent in one-on-one research mentor–trainee relationships by providing complementary guidance and structure (Chesler and Lohman, 1971; Darling, 1986; Head *et al.*, 1992; Hurtado *et al.*, 1998, 2008; Kim *et al.*, 1998; Bauer and Bennett, 2003; Eby *et al.*, 2004; Fenning, 2004; Mullen, 2007; Wisker *et al.*, 2007; Anderson and Shore, 2008; Chan, 2008; Byars-Winston *et al.*, 2011; Bartlett, 2012; Ngassa, 2013). The support provided by structured research training programs can be particularly valuable for trainees from underrepresented backgrounds (Carter *et al.*, 2009; Packard, 2015) and can therefore advance efforts to diversify the research workforce. Structured research training programs help mentors, research teams, and thesis committees understand the challenges that trainees are facing, particularly challenges that emerge from the intersection of a trainee's unique racial or cultural background and the traditional research training culture (Museus and Liverman, 2010; Packard, 2015; Godwin *et al.*, 2016; Layton *et al.*, 2016; Carver *et al.*, 2017; Research Triangle International, 2018). Without program support, trainees whose experiences and values do not align with traditional research culture may struggle to develop the social and cultural capital they need to successfully navigate and persist in research (Bauer and Bennett, 2003; Hurtado *et al.*, 2008; Lewis *et al.*, 2016; Mau, 2016; Carver *et al.*, 2017). The *Entering Research* curriculum is a collection of activities and resources that develop the skills and knowledge undergraduate and graduate research trainees need to navigate the research culture and succeed in research.

Entering Research: A Curriculum for Structured Research Training Programs

Entering Research, 1st Edition. The first edition of the *Entering Research* (ER) curriculum (Branchaw *et al.*, 2010) was developed to optimize the undergraduate apprentice-style research learning experience by providing just-in-time support and community. The first edition consisted of a two-semester, student-centered, active-learning curriculum designed for undergraduate research training programs and courses. It created community, structured the research experience, clarified behavioral and performance expectations, and provided information and resources to develop the social and cultural capital that beginning undergraduate researchers need to successfully navigate the research environment. The first edition addressed many of the recommendations put forth in two recent studies from the National Academies of Sciences, Engineering, and Medicine on undergraduate research (NASEM, 2017) and graduate education (NASEM, 2018), which document the importance of research learning experiences at undergraduate and graduate career stages, highlight the potential of research learning experiences to advance efforts to diversify the STEMM research workforce, and call for student-centered, evidence-based approaches to be used in designing research training programs. Research on the first edition showed evidence of increased learning outcomes for student researchers who participated in an ER course relative to a comparison group of research students who did not partici-

pate in the course at the same institution (Balster *et al.*, 2010). Students who participated in an *Entering Research* course reported statistically significantly higher gains in their skills, knowledge, and confidence as researchers relative to a comparison group of students engaged in undergraduate research at the same institution, but not enrolled in the course.

Entering Research, 2nd Edition. The evidence of effectiveness with undergraduate research trainees and positive feedback from users of the first edition of the ER curriculum motivated development of the second edition of the curriculum, which was funded by the National Institute of Health's (NIH) Diversity Program Consortium's National Research Mentoring Network (NRMN; www.diversityprogramconsortium.org). The second edition was significantly reorganized, expanded, and adapted for use with graduate students as well as undergraduate students (Branchaw *et al.*, 2020). A multidisciplinary team of natural and social scientists with scholarly expertise and practical experience working with diverse research trainee populations at a variety of types of institutions was convened to adapt existing and develop new curricular activities for the second edition (see Supplemental Material for a complete list of development team members and institutions). An iterative process of adaptation/development, review, and revision was used to deconstruct and expand the original two-semester curriculum into individual activities and to develop new activities. A common template and common rubric¹ were used during development to set internal review criteria and to establish consistency of content. The second edition has 96 individual activities from which research training program directors can select to construct custom curricula for use with small cohorts of 10–15 undergraduate or graduate students. Individual activities can be integrated into existing research training program workshops or courses, or new workshops and courses can be created.

Activities to build research trainees' awareness and skills to manage equity and inclusion issues were developed for the second edition to specifically address equity and inclusion in the STEMM research environment. In addition, "inclusion considerations" were added to the implementation guide for every activity to support research training program directors to adapt activities for use with trainees of different backgrounds and levels of preparation. A conceptual framework with seven areas of trainee development and learning objectives was created to organize the 96 activities in the second edition based on factors identified in "foundational research" studies (Institute of Education Sciences and the National Science Foundation [IES and NSF], 2013) as important for research trainee development (Table 1; a complete list of activities is in the Supplemental Material). Evidence of validity of the hypothesized structure of the framework was collected during development of the *Entering Research* Learning Assessment (ERLA; Branchaw and Butz, 2019; Butz and Branchaw, 2020, in press).

In this paper, we report results from two "design and development research" studies investigating 1) the effectiveness of the second edition of the ER curriculum in a nationwide pilot test and 2) the value of a workshop to train research training program directors (facilitators) to use the second edition curriculum. The *Common Guidelines for Education Research*

¹Activity templates and rubrics are available upon request from the authors.

TABLE 1. Entering Research conceptual framework and example activities^a

Areas of trainee development with learning objectives	Example activities
Research Comprehension and Communication Skills (RCC; 36% of curriculum)	
<ul style="list-style-type: none"> • Develop Effective Interpersonal Communication Skills • Develop Disciplinary Knowledge • Develop Research Communication Skills • Develop Logical/Critical Thinking Skills • Develop an Understanding of the Research Environment 	<ul style="list-style-type: none"> ○ Addressing Conflict ○ Aligning Mentor and Trainee Expectations ○ Prioritizing Research Mentor Roles ○ Your Research Group's Focus ○ Communicating Research Findings 3: Developing Your Presentation ○ Research Writing 1: Background Information and Hypothesis or Research Question ○ Research Writing 7: Research Paper
Practical Research Skills (PRS; 6% of curriculum)	
<ul style="list-style-type: none"> • Develop Ability to Design a Research Project • Develop Ability to Conduct a Research Project 	<ul style="list-style-type: none"> ○ Research Writing 3: Project Design ○ Searching Online Databases ○ Safety Training Checklist
Research Ethics (RE; 9% of curriculum)	
<ul style="list-style-type: none"> • Develop Responsible and Ethical Research Practices 	<ul style="list-style-type: none"> ○ Case Study: Authorship ○ Truth and Consequences Article ○ Research Writing 5: Peer Review Process
Researcher Identity (RID; 6% of curriculum)	
<ul style="list-style-type: none"> • Develop Identity as a Researcher 	<ul style="list-style-type: none"> ○ Networking 3: Your Brand ○ Personal Statement ○ Developing a Curriculum Vitae
Researcher Confidence and Independence (RCI; 7% of curriculum)	
<ul style="list-style-type: none"> • Develop Confidence as a Researcher • Develop Independence as a Researcher 	<ul style="list-style-type: none"> ○ Steps to Researcher Independence ○ Fostering Your Own Research Self-Efficacy ○ Case Study: Overwhelmed
Equity and Inclusion Awareness and Skills (EI; 11% of curriculum)	
<ul style="list-style-type: none"> • Develop Skills to Deal with Personal Differences in the Research Environment • Advance Equity and Inclusion in the Research Environment 	<ul style="list-style-type: none"> ○ Challenges Facing Diverse Teams ○ Privilege and White Fragility ○ Counter-Storytelling ○ Stereotype Threat
Professional and Career Development Skills (PD; 21% of curriculum)	
<ul style="list-style-type: none"> • Explore and Pursue a Research Career • Develop Confidence in Pursuing a Research Career 	<ul style="list-style-type: none"> ○ The Next Steps in Your Career ○ My Mentoring and Support Network ○ Research Careers: Informational Interview ○ Letter of Recommendation

^aPercentages reflect the proportion of activities that primarily fall within each area of trainee development relative to the total number of activities. A full listing of all activities available in the *Entering Research* curriculum is available in the Supplemental Material. Activities can also be searched by area of trainee development on the CIMER website, <http://cimerproject.org>.

and *Development Report* (IES and NSF, 2013) defines design and development research as research to collect initial evidence on an intervention that can be used as the foundation for more advanced types of “efficacy, effectiveness, and scale-up research.” The results of the two studies presented here provide evidence that the second edition of the ER curriculum is ready for use by undergraduate and graduate research training program directors, that it is sufficiently promising to warrant future efficacy, effectiveness, and scale-up research, and that the ER training workshop is an effective way to train research program directors to use the curriculum.

STUDY 1 (S1)

Design and Development Research on the 2nd Edition of Entering Research

The first study sought to determine whether the second edition of the ER curriculum achieved the intended outcomes under

various conditions through pilot testing with research training program directors (facilitators) and undergraduate and graduate students (research trainees) across the country. The primary goal of pilot testing the second edition of the ER curriculum was to measure whether and to what extent the facilitators and research trainees found the curricular activities to be complete, clear, and valuable. The following evaluation questions (EQs) guided our study:

EQ1.1: How is the ER curriculum implemented?

EQ1.2: To what extent are the materials provided to facilitators useful in their current form?

EQ1.3: How effective are the ER activities in helping trainees gain knowledge and/or improve their ability to do research?

EQ1.4: What is the perceived value by trainees of participating in ER activities?

EQ1.5: What are trainees' learning gains after participating in an ER implementation?

EQ1.6: To what extent are gains in research knowledge or ability attributed by trainees to participating in ER activities versus the research experience?

EQ1.7: What is the impact of ER implementations on trainees' interest in conducting research in the future?

S1 METHODS

Participants

Sites were recruited for pilot testing through the NRMN (<https://nrmnet.net>) and through the ER facilitator training workshops. Overall, there were 42 unique pilot implementations: 31 implementations with undergraduate student participants, 10 implementations with graduate student participants, and one implementation with a combination of undergraduate and graduate student participants. Implementations occurred at 20 different sites (75% doctoral-granting institutions; 10% master's-granting colleges; 10% medical colleges; 5% nonprofit organizations; 15% minority-serving institutions). Trainees and facilitators were surveyed at each site. Trainees from three implementations and facilitators from two implementations did not complete evaluation surveys; overall, data from 591 research trainees and 78 facilitators were collected. Of the 591 trainees who completed at least one evaluation survey (79% average response rate), 570 provided consent for their evaluation survey responses to be used in research (Institutional Review Board [IRB] protocol 2017-0026). Data from five participants who reported their age as under 18 were removed from the data set, resulting in a final sample size of 565 trainees. Demographic information about the study participants is in Table 2, and copies of the pilot testing evaluation surveys are in the Supplemental Material.

Procedure

In pilot testing, activities were either 1) integrated into existing courses or programs alongside non-ER activities or 2) new courses, programs, or stand-alone workshops were developed that exclusively used ER activities. When selecting activities for implementation, pilot testing facilitators were encouraged first to identify the learning objectives for their trainees and then to use the ER conceptual framework and their trainees' career stage (undergraduate or graduate) and level (novice, intermediate, or advanced) to identify appropriate activities. This backward design process for curricular development (Wiggins *et al.*, 1998) was either conducted via telephone consultations with an author of the curriculum or was part of the facilitator training workshop in which some of the pilot testing facilitators participated. Subsequently, this process has been codified in the published curriculum (Branchaw *et al.*, 2020), where a step-by-step guide for creating custom curricula is available.

Feedback from research trainees and facilitators was collected via online surveys that were deployed either at the end of a workshop series or, in the case of classes or programs lasting more than a few weeks, deployed at various time points throughout the program to ensure participation in the ER activities was fresh in the minds of participants (IRB protocol 2016-0458). Consequently, some questions had lower response rates (and therefore lower sample sizes), because they were asked on a later survey. The timing of these surveys varied

based on facilitator preference and the scheduling of ER activities within a given course or program.

Data Sources and Analyses

Recording Implementations of ER (EQ1.1). Facilitators were asked to forward copies of their syllabi or program schedules before pilot testing so the research team could build custom evaluation surveys asking about specific activities. After implementation, but before surveys were deployed, facilitators confirmed which activities had actually been implemented. The percentage of activities implemented from each area of trainee development was calculated by taking the number of activities implemented by a facilitator from a given area of trainee development and dividing it by the total number of activities implemented. The types of programs in which activities were implemented (e.g., summer programs, courses, workshops) were also tracked.

Evaluating Ease of Implementing ER (EQ1.2). To understand the extent to which the format and content of the ER activities were useful to facilitators, we asked them to share whether they had enough time to facilitate activities, whether the facilitator notes were clear and complete, and whether the student materials and handouts were clear and complete. Facilitators answered each question with “yes” or “no” and were given the option to provide specific comments through open-ended prompts. The frequency and percentage of “yes” responses to questions about enough time to facilitate an activity, the clarity and completeness of the facilitator notes, and the clarity and completeness of the student materials were calculated to gauge facilitator satisfaction with and perceived ease of use. Open-ended responses were also reviewed to identify key themes across implementations and were used formatively to guide ongoing activity refinement.

Evaluating ER Activity Effectiveness (EQ1.3). Overall activity effectiveness was measured by asking trainees and facilitators: “From your perspective, please rate how effective this activity was in helping you (your trainees) gain knowledge and/or improve your (their) ability to do research?” A five-point response scale was provided ranging from 1 (very ineffective) to 5 (very effective). The average effectiveness of activities in each primary area of trainee development was measured by calculating the average effectiveness of each individual activity, then calculating an average effectiveness of activities under each area of trainee development.

Evaluating Perceived Value of ER Curriculum (EQ1.4). Trainees were asked three questions to assess their satisfaction with and the perceived value of ER activities. 1) “Overall, how effective were the facilitators in guiding discussion during your ER activities?” 1 (very ineffective) to 5 (very effective); 2) “How likely are you to recommend participation in ER activities?” 1 (very unlikely) to 5 (very likely); and 3) “Overall, was participation in the ER activities a valuable use of your time?” “yes” or “no.” The average rating of facilitator effectiveness, the average likelihood of recommending participation in ER activities, and the percentage of individuals who responded “yes” when asked whether participation in ER activities was a valuable use of their time were used to measure trainees' perceived value of the ER curriculum.

TABLE 2. Trainee demographic information (N = 565)^a

Gender	Percent
Female	50.1
Male	29.7
Other gender identity	<1
Not reported	20.0
Race/ethnicity	
Native American/Alaskan Native	1.2
White	43.4
Asian	10.4
African American	9.7
Multiple races	5.0
Other	5.5
Not reported	24.8
Hispanic	21.8
Training stage	
First- or second-year undergraduate student	22.7
Third-, fourth, or fifth-year undergraduate student	46.4
Postbaccalaureate student	4.2
First- or second-year graduate student	13.8
Third-, fourth, or fifth-year graduate student	2.8
Other	<1
Not reported	9.4
Prior research experience	
Yes	55.8
No	34.5
Not reported	9.7

^aRespondents could select Hispanic in addition to a race category. As a result, total percentages for the sample may add up to more than 100%.

Evaluating the Impact of ER Implementations on Trainee Learning (EQ1.5 and EQ1.6). The ERLA was administered to assess trainee learning gains and the impact of ER implementations on trainee learning (EQ 1.5). The ERLA assesses trainee learning gains in the seven areas of trainee development articulated in the ER conceptual framework (Branchaw and Butz, 2019; Branchaw *et al.*, 2020; Butz and Branchaw, 2020, in press). Validity evidence for test content, internal structure, convergence, and evidence of internal consistency of the ERLA was collected and assessed using two separate samples of trainees and mentors and was found to have acceptable model–data fit and internal consistency statistics greater than 0.81 for each subscale (Butz and Branchaw, 2020, in press). The ERLA consists of 53 items, and responses range from 1 (no gain) to 5 (great gain). Subscale means for each area of trainee development were calculated from trainees' responses. A subset of facilitators opted to administer the ERLA as part of their evaluation surveys; 111 trainees completed the ERLA. Average learning gains reported for each area of trainee development were calculated based on trainees' responses to the ERLA items corresponding to each area of trainee development in the conceptual framework.

Trainees in 37 implementations were asked to attribute their gains in research knowledge or ability to participating in the ER activities; trainees in 34 implementations were asked this same question in relation to their research experiences (EQ 1.6). Responses could range from 1 (none) to 5 (a great deal). The frequency with which trainees attributed a *fair amount* or a

great deal of their learning gains to the ER activities was used to measure the impact of participation in the activities compared with the impact of doing research.

Evaluating the Impact of ER Implementations on Trainees' Interest in Research (EQ1.7). Trainees were asked whether their experiences in the ER course/seminar/workshop increased their interest in conducting research in the future, with response options ranging from 1 (strongly disagree) to 5 (strongly agree), and were provided a text box to explain their answers. A mean response to the item was calculated, and open-ended comments associated with the responses indicating agree or strongly agree ($n = 196$) and disagree or strongly disagree ($n = 15$) were coded to identify key themes raised by participants (EQ1.7). Responses were open-coded using a multistep process (Creswell, 2009) to allow themes to emerge. First, all open-ended responses were reviewed, and key themes common across many responses were identified by a researcher (A.R.B.). Next, all responses were coded using the themes. During coding, some themes were combined and reorganized. The results were reviewed by a second researcher (J.L.B.), and any discrepancies were discussed until agreement was reached. As a result of this discussion, the initial themes identified were re-examined, and additional codes were added to identify the extent to which trainee responses aligned with the areas of trainee development outlined in the ER conceptual framework. Final codes were counted across all themes to determine those most frequently identified. The final codes, along with definitions and example quotes, are presented in the *S1 Results and Discussion*.

S1 RESULTS AND DISCUSSION

EQ1.1: Activities from the ER Curriculum Were Successfully Implemented in a Number of Ways

To evaluate how research training program directors would use the activities, we invited pilot testers to select and test activities that would meet their trainees' needs. Consequently, some of the activities were tested more extensively than others. In total, 85 activities were implemented by facilitators; 70 of those activities (88%) were implemented at least twice. The total number of ER activities included in each implementation ranged from one to 27; the average implementation included nine activities. There were 10 activities for which no facilitator or trainee data were available. Seven were new activities and five were from the Research Comprehension and Communications Skills area of trainee development. Four of the five addressed thesis writing at the undergraduate or graduate levels, which made them difficult to implement as part of a workshop or course lasting one semester or less.

Overall, the percentage of activities in each area of trainee development used during pilot testing generally reflected the percentage of those activities in the curriculum. The highest percentage (42%) of activities implemented were from the Research Comprehension and Communication Skills area, which makes up 36% of the activities. Their overrepresentation in implementations suggests that these activities align well with the goals of existing research training programs and courses. By contrast, the equity and inclusion awareness and skills activities represent 11% of the curricular activities, but accounted for only 5% of the activities implemented by pilot testers. We learned through consultations and the discussions in facilitator

training workshops that many of the pilot testers were less confident of their ability to implement these activities compared with the other activities. Consequently, even though pilot testers and workshop participants indicated that they thought equity and inclusion activities were very important, they implemented these activities less frequently. This is not surprising, given that most research training programs do not prepare scientists to address equity and inclusion (Davidson and Foster-Johnson, 2001; Prunuske *et al.*, 2013; Butz *et al.*, 2018). Some facilitators have indicated they would prefer to give the responsibility for facilitating these activities to outside experts, but we have discouraged them from doing this, because it could suggest to trainees that advancing equity and inclusion is not the responsibility of scientists. Instead, we encouraged facilitators to invite outside experts to cofacilitate these activities with them to build their knowledge and confidence to assume full responsibility for facilitating the activities in the future.

Of the 42 implementations included in pilot testing, 11 (26%) incorporated ER activities as part of a Summer Undergraduate Research Experience, nine (21%) as part of a workshop or workshop series, two (5%) as part of a seminar, and 20 (48%) as part of a course. The diversity of types of implementations represented in this data provide evidence that the curriculum is flexible and accessible to research training program directors working in a variety of settings.

EQ1.2: Facilitators Reported the Implementation Guides and Trainee Materials to Be Useful and Were Able to Modify Activities to Suit Their Needs

Overall, facilitators were able to use the ER curricular materials successfully. Across the 40 unique implementations for which we received facilitator surveys, 99% of facilitators answered “yes” when asked if the facilitator notes and trainee materials were clear and complete for the activities they implemented. This suggests that the ER activities were perceived by facilitators to be of consistent quality and that the rigorous, iterative review process used to develop activities yielded consistently complete facilitator notes and trainee materials.

The positive numerical results were echoed in the comments that facilitators provided. For example, one facilitator shared, “The notes are very good. They provided guidance and talking points that were helpful in getting students engaged.” Nonetheless, there was a range of responses, with a few facilitators commenting that the notes were too detailed, and others asking for additional support and resources. We were not able to further analyze these comments based on the relative experience or self-perceived competence of facilitators, but responses indicating that there was too much or too little information suggest that various levels of detail in the implementation guides are needed.

Facilitators shared similar positive feedback about the trainee materials, noting that most materials provided clear expectations and that they felt comfortable modifying the materials to suit their trainees’ needs and the contexts of their implementations. For example, one facilitator noted, “I felt the student materials were also clear and complete. More importantly, the students found them clear and complete. No one seemed confused by them.” A few facilitators noted that it was difficult to find the supporting materials referenced in the trainee handouts. We have since integrated these supporting materials into the activities whenever possible.

Facilitators reported that they had enough time to facilitate the activity for 94% of the activities that were implemented. When activities ran over time, facilitators frequently commented that it was due to ongoing discussions that they did not want to end. For example, one facilitator shared, “We ran out of time on Elevator Sentences; we took our leisure to understand and give comprehensive feedback for each participant. I am glad we did that despite using at least double the amount of time initially allotted to that activity.” When activities ran under time, facilitators often attributed it to the size or characteristics of the group or to modifications that they had made. One facilitator noted, “I found that the time allocated for some of the activities could be shortened based on the level of exposure of the learners. I had a smaller group, so sometimes it wasn’t worth doing a think-pair-share.” At times, facilitators noted that the constraints of the implementation (e.g., having a set amount of time for each session) led them to modify the timing of activities. We interpret these responses as an indication that facilitators were able to modify the activities to suit their programs’ needs.

A criticism of the first edition of the ER curriculum was that the two-semester format was rigid and made it difficult for research training program directors and instructors to identify and modify activities for use in their programs. These evaluation data suggest that this limitation has been addressed with the reorganization of the second edition and that facilitators are able to effectively identify and implement relevant activities using the guidelines provided. Importantly, the data also reveal that the structure of the activities gave facilitators permission and confidence to modify them without too much effort, which we encourage facilitators to do when integrating ER activities into existing programs or courses.

EQ1.3: Reported Effectiveness of ER Activities in Helping Trainees Gain Knowledge and/or Improve Their Ability to Do Research

For reporting purposes, the trainee and facilitator individual activity effectiveness data are grouped by the primary activity area of trainee development (Figure 1). Overall, the activities were rated as effective by both facilitators and trainees, with slightly higher ratings from facilitators. Facilitators rated activities from the Research Comprehension and Communication Skills area of trainee development highest ($M = 4.39$; $SE = 0.090$) and activities from the Research Ethics area of trainee development lowest ($M = 4.17$; $SE = 0.063$). Trainees rated activities from the Professional and Career Development Skills area of trainee development the highest ($M = 4.00$; $SE = 0.047$) and activities from Equity and Inclusion Awareness and Skills the lowest ($M = 3.82$; $SE = 0.092$).

Though the overall variability of facilitator and trainee ratings across the areas of trainee development is relatively small (from 3.82 to 4.39), these data show a mismatch between what research training program directors perceive to have had the greatest impact on their trainees’ development (Research Comprehension and Communication Skills activities) and what their trainees perceived to have had the greatest impact on their development (Professional and Career Development activities). Such a mismatch suggests that research training program directors should gather input from their trainees about their needs before selecting activities and that they should monitor their

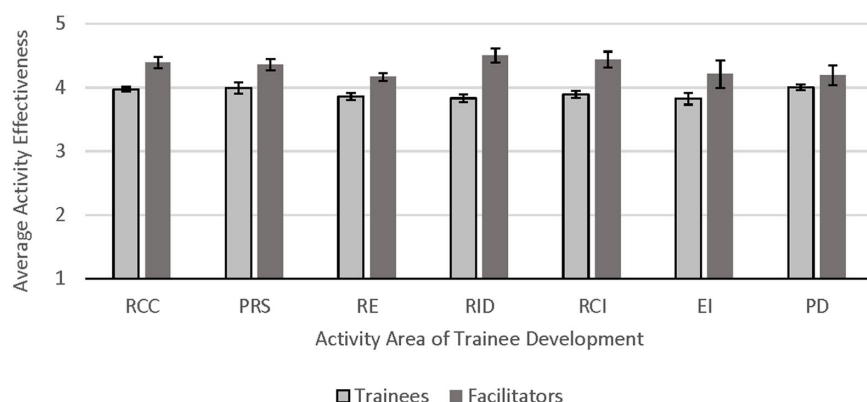


FIGURE 1. Average activity effectiveness by activity primary area of trainee development trainee prompt: From your perspective, please rate how effective this activity was in helping you gain knowledge and/or improve your ability to do research? Facilitator prompt: From your perspective, how effective was the activity in helping mentees gain knowledge and/or improve their ability to do research? Responses could range from 1 (very ineffective) to 5 (very effective). Error bars reflect the SEM. RCC, Research Comprehension and Communication Skills; PRS, Practical Research Skills; RE, Research Ethics; RID, Researcher Identity; RCI, Researcher Confidence and Independence; EI, Equity and Inclusion Awareness and Skills; and PDS, Professional and Career Development Skills. Trainee and facilitator per unique implementation N : RCC = 425, 34; PRS = 108, 10; RE = 207, 16; RID = 258, 21; RCI = 251, 24; EIA = 134, 11; PD = 255, 28.

trainees' perceptions and learning throughout implementation to determine whether real-time curricular adjustments should be made.

As mentioned earlier, facilitators were generally less comfortable implementing the Equity and Inclusion Awareness and Skills activities, and fewer implementations of these activities were reported. The small number of implementations coupled with the fact that these activities were probably less effectively facilitated based on our pilot testers' confidence levels may account for the lower ratings from trainees. In addition, the common perception that equity and inclusion are not core to the STEMM research mission (Prunuske *et al.* 2013; Butz *et al.*, 2018) may have led to dissatisfaction with the inclusion of these activities in a program or course traditionally focused on developing research skills and knowledge and preparing future researchers for successful STEMM careers. We argue, however, that only by integrating these types of activities into research training curricula and engaging established and aspiring STEMM researchers in conversations about equity and inclusion, even if they feel uncomfortable, will STEMM research culture become more inclusive and the discussion of equity and inclusion become increasingly perceived as central to trainee development.

EQ1.4: Trainees Valued Their Participation in ER Implementations

Overall, trainees ($n = 304$) rated facilitators as effective in guiding ER activity discussions, with an average rating of 4.41 on a scale ranging from 1 (very ineffective) to 5 (very effective; $SE = 0.044$), 80% ($n = 242$ of 303) reported that they were likely or very likely to recommend participation to their peers, and 90% reported that participation in the ER activities was a valuable use of their time ($n = 274$ of 303). Together, these results indi-

cate that facilitators were able to effectively engage trainees in the ER activities and that trainees valued the experience.

EQ1.5: Trainees Participating in ER Activities Reported Learning Gains across All Areas of Trainee Development

A subset of pilot testing trainees ($n = 111$) completed the ERLA during Spring and Summer 2018 implementations. Gains in all areas of trainee development were above 3.80 on a scale ranging from 1 (no gain) to 5 (great gain). Trainees reported learning gains in all areas (Figure 2), with the largest gains in Research Comprehension and Communication Skills (4.17; $SE = 0.065$) and the smallest gains in Equity and Inclusion Awareness and Skills (3.80; $SE = 0.091$), reflecting the most and least addressed areas of trainee development, respectively.

The alignment of learning gains with the number of ER activities implemented in each area of trainee development suggests that implementations with more activities addressing a particular area of trainee development yield greater learning

gains in that area. However, without a comparison group or data to control for other non-ER activities and the research experience itself, it is impossible to establish a causal effect. Future research will test this. Regardless, practitioners interested in implementing ER activities should assess student learning using multiple measures. For example, in addition to measuring student self-assessments of learning gains using the ERLA, the matched ERLA mentor survey can be used to measure and compare the mentor's assessments of the trainee's learning gains. In addition, assessment of research products that reflect student learning gains such as papers and presentations can be made using common rubrics. Several rubrics are included in the ER curriculum.

EQ 1.6: Research Experiences and Participation in ER Activities Both Contribute to Trainee Perceived Gains in Research Knowledge or Ability

Beyond asking trainees to rate the activities and self-assess their learning gains in the various areas of trainee development, we also began to explore the extent to which participation in ER activities contributed to trainee gains relative to the gains they experienced from doing research alone by asking them to attribute their gains in research knowledge and ability to their participation in ER activities or to doing research (Figure 3). As expected, students attributed a fair amount (24%) or a great deal (63%) of their gains in research knowledge and ability to their participation in the research experience. However, many also reported that participating in a course or workshop with ER activities contributed a fair amount (36%) or a great deal (31%) to their gains in research knowledge and ability. This suggests that activities in the areas of trainee development focused on developing these skills, namely Research Comprehension and Communication Skills and Practical Research

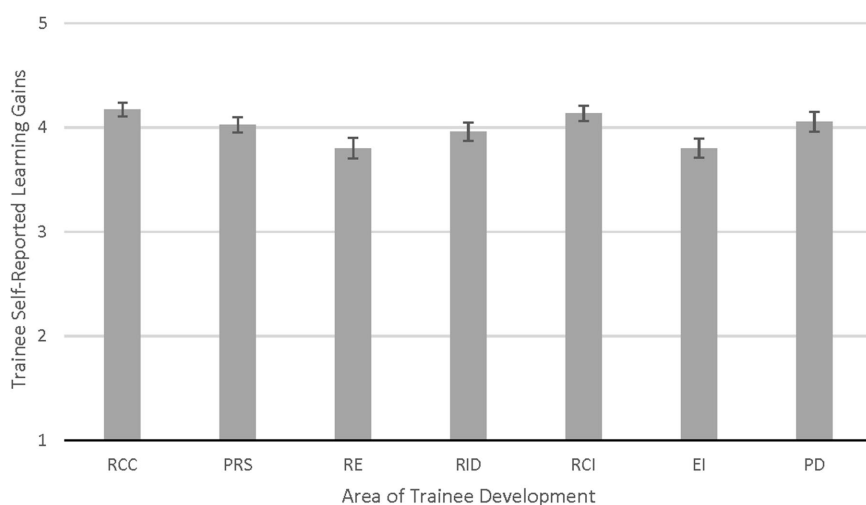


FIGURE 2. Average learning gains reported by trainees who participated in ER interventions. Responses could range from 1 (no gain) to 5 (great gain). Each mean corresponds to an area of trainee development addressed in the ER curriculum. Error bars represent the SEM. RCC, Research Comprehension and Communication Skills; PRS, Practical Research Skills; RE, Research Ethics; RID, Researcher Identity; RCI, Researcher Confidence and Independence; EI, Equity and Inclusion Awareness and Skills; and PDS, Professional and Career Development Skills. $N = 112$.

Skills, provided complementary and useful support to trainees engaged in doing research.

To explore the value added by ER activities further, we will incorporate data from comparison groups and learning attribution questions for all areas of trainee development into future research on the effectiveness and impact of the curriculum. Even so, because the ER curriculum is process based and the activities are designed to be implemented while trainees are engaged in the research learning experience, we anticipate that it will be challenging to disentangle and attribute learning gains resulting from the research experience from learning gains resulting from participation in ER activities. Consequently, those interested in implementing ER activities should consider not only which activities will address gaps in training in their programs, but how and when those activities will provide the just-in-time support their trainees need to be successful doing research. For

example, activities that help trainees to establish positive relationships with their mentors and research group colleagues will likely have the greatest impact at the beginning of the research experience, while activities that develop trainees' research writing and presentation skills will likely have the greatest impact when trainees are preparing to submit their research for publication or to present at a conference.

EQ1.7: Experiences in ER Implementations Contribute to Trainees' Reported Interest in Conducting Research in the Future in a Number of Ways

Of the 273 trainees who responded to the question "My experience in this course/workshop increased my interest in conducting research in the future," 71% ($n = 194$) agreed or strongly agreed. The average response was a 3.94 ($SE = 0.055$) on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). Codes and definitions that were assigned to open-ended responses and the frequency with which each code was assigned are reported in Table 3. Of the

194 trainees who indicated that they agreed or strongly agreed that their participation in the course/workshop increased their interest in conducting research in the future, 146 (75%) provided a response to the open-ended prompt. Sixteen (8%) individuals indicated that they strongly disagreed or disagreed with this statement, and of those, 11 provided a response to the open-ended prompt.

Trainees reported that experiences or skill gains related to Professional and Career Development Skills (36%) and Research Comprehension and Communication Skills (32%) had the greatest impact on their interest in future research experiences and training. Trainees who reported that a course or program that included ER activities helped them clarify their research or career interests and plans often commented that it allowed them to see a possible path (or multiple pathways) to a research career, or, alternatively, how a career in research was not the right path for them, indicating that ER activities prompt trainees to think about whether a career and life in research is the right fit for them. Trainees who noted that their interest increased because of opportunities to build or apply Research Comprehension and Communication Skills reported that the ER activities helped to demystify the research experience and the norms and expectations of research.

The open-ended responses also provided insight into how specific ER activities either complemented or filled gaps in trainee research learning experiences (27%) and provided direct references to the research experience or to research mentors; 13% of trainee responses noted that the research experience itself

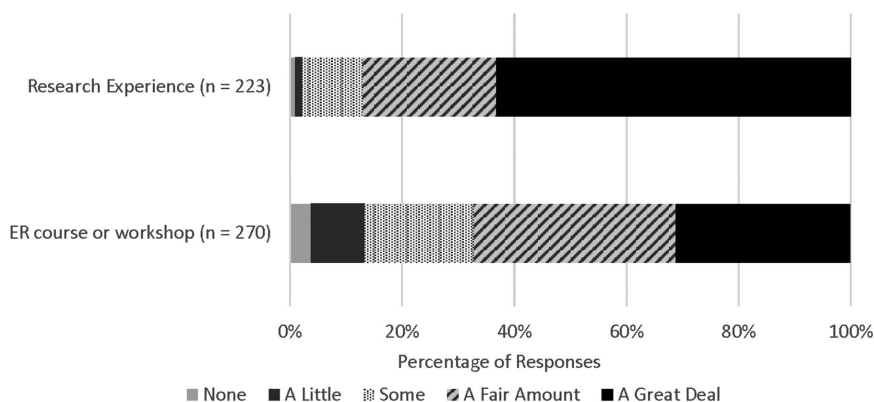


FIGURE 3. Trainee learning attribution. Trainees were asked to report how much their experience doing research and their experience participating in ER activities contributed to their gains in research knowledge or ability. $N = 225$.

TABLE 3. Codes and example responses of reasons why trainees' interest in research increased or decreased

Code and definition	Example responses	Number of times code assigned ^a	
		Increased Interest (N = 146)	Decreased Interest (N = 11)
Research Experience. Describes impact or role of the research experience or research mentor.	<ul style="list-style-type: none"> <i>Increased:</i> "Having a great experience in my lab with a mentor who was there for guidance increased my interest in conducting research in the future" <i>Decreased:</i> "My graduate student mentor was not helpful and made me feel not welcomed and uncomfortable. I always felt as an inconvenience to him." 	19	2
Skills and Experience —Describes experience or skill gain related to each area of trainee development			
Professional and Career Development Skills	<ul style="list-style-type: none"> <i>Increased:</i> "I think understanding that I can really translate skills to any area of research was useful. It made me further understand that I will not be 'stuck' conducting the same research my entire life." <i>Decreased:</i> "I gained a lot of respect for the people that pursue a career working in research, but I don't believe that it would be something that I would pursue thanks to this experience." 	53	4
Research Comprehension and Communication Skills	<ul style="list-style-type: none"> <i>Increased:</i> "Being able to understand that there is more to research than just results. That it is about interacting with different people to try and reach some conclusion, that will lead us to another topic. That idea is what stuck with me from these discussions." <i>Decreased:</i> "This workshop enforced that science is a hierarchical field that is difficult to break into." 	47	1
Researcher Confidence and Independence	<ul style="list-style-type: none"> <i>Increased:</i> "Increased the idea of possibilities and combating ideas of imposter syndrome and lack of self-worth. Overall the experience is quite energizing and stimulating." 	18	0
Practical Research Skills	<ul style="list-style-type: none"> <i>Increased:</i> "Having the knowledge and tools to conduct even basic research helps me feel less anxious about pursuing further research opportunities." 	13	0
Equity and Inclusion Awareness and Skills	<ul style="list-style-type: none"> <i>Increased:</i> "Being able to reframe a negative experience makes you better equipped to face challenges in the lab in the future. It allows you to work through that hard moment and be ready for future hardships." 	5	0
Research Ethics	<ul style="list-style-type: none"> <i>Increased:</i> "I feel my interest has increased; however, the class has brought to light the reality of working in research including the competitiveness and the importance of ethics and transparency in research." 	4	0
Researcher Identity	<ul style="list-style-type: none"> <i>Increased:</i> "Having the opportunity to conduct research in a different university and laboratory made me realize that no matter where I am at, I can see myself having a career as a research scientist." 	2	0
Workshop, Course, or Activity. Reference to a course or workshop featuring activities covered in the ER curriculum or specific activities from the ER curriculum.	<ul style="list-style-type: none"> <i>Increased:</i> "Methods taught to aid better literature reading and scientific writing helped a lot." <i>Decreased:</i> "I do not enjoy the class but I do enjoy the research aspect itself. My mentor ... is a wonderful professor and leader." 	39	4
Entering Research Cohort. Describes the impact of the community created by ER implementations.	<ul style="list-style-type: none"> <i>Increased:</i> "From the workshops, it showed me that other people have similar thoughts about their research experience (good or bad) and it is ok to talk about these situations without being embarrassed. This provided the feeling of a support network in research and encouraged me to keep going." 	12	0
Facilitator. Reference to a facilitator.	<ul style="list-style-type: none"> <i>Increased:</i> "The enthusiasm/knowledge of the instructor." 	2	0
General Interest, Experience or Skills. Trainee expressed that interest increased in general, or referenced their experience, but did not provide specific examples.	<ul style="list-style-type: none"> <i>Increased:</i> "Overall I learned so much about research and I have really liked it so far. I feel really good about the skills I have gained." <i>Decreased:</i> "Nothing could have increased my interest, I am maximally interested in research." 	13	1
Other. Responses that did not directly address the question.	<ul style="list-style-type: none"> <i>Increased:</i> "I have taken many courses that have increased my interest in research such as, cell physiology, biochemistry, and genetics." 	2	0

^aResponses could be assigned multiple codes, so total number of codes may exceed the total number of responses.

increased their interest, and 1% said that it decreased their interest. One response in particular reflects how research experiences and a program that incorporates ER activities can profoundly impact a trainee's interest in continuing in research: "Having the opportunity to conduct research in a different university and laboratory made me realize that no matter where I am at; I can see myself having a career as a research scientist. Additionally, the workshops/seminar provided made me realize that there will of course be challenges faced in this field, especially as a minority in STEM, however, learning about the ways that can help me through it gave me confidence to continue pursuing this career."

Overall, the open-ended responses provide insight about which ER activities may have the greatest influence on research trainees and how the activities are valued, or not, by trainees. In addition, all areas of trainee development in the ER conceptual framework were represented in the comments, further confirming that the ER conceptual framework captures the various dimensions of research trainee development.

S1 CONCLUSIONS

The first design and development research study provides evidence that research training program directors are willing and able to use the second edition of the ER curriculum and that participation in the activities is valued by research trainees. Pilot testing data indicate that the activity template, review rubric, and iterative development process yielded well-constructed accessible activities and that the revision and expansion of the curriculum did not negatively affect the gains in research trainee skills, knowledge, and confidence reported in the study on the first edition of the curriculum (Balster *et al.*, 2010). Moving forward, this gives us confidence that facilitators will be able to implement the ER activities with fidelity and that trainees will fully participate in them in future efficacy, effectiveness, and scale-up research studies investigating the long-term impacts of the second edition of the ER curriculum on research trainee learning gains, development, and career trajectories.

STUDY 2 (S2)

Design and Development Research on an *Entering Research Facilitator Training Workshop*

The second study sought to determine whether a two-day ER facilitator training workshop effectively prepared research training program directors to design and implement training interventions using activities from the second edition of the ER curriculum. The ER facilitator workshop employs the backward design process (Wiggins *et al.*, 1998) to train facilitators to first define learning objectives for their research trainees and then to use the ER conceptual framework to identify activities and aligned evaluation and assessment tools to use in their research training program intervention. The workshop also provides training on best practices in facilitation and implementation planning.

The first day of the workshop includes an introduction to the ER conceptual framework, activities, and assessment and evaluation tools; an introduction to backward design; instruction on how to use the ER conceptual framework to sort and find activities; and an opportunity to practice facilitating ER activities, all of which rely on active-learning pedagogies. On the second day of the workshop, participants are provided structured time to work on developing their ER training interventions.

They each build a custom ER curriculum and develop implementation and assessment/evaluation plans. Participants share and discuss drafts of their trainee learning objectives, the ER activities they have selected to address those objectives, the learning assessment and program evaluation instruments they plan to use, and their implementation plans. The second day concludes with a structured discussion of the challenges they may face in facilitating activities, assessing student learning, and implementing and evaluating their intervention implementations, with a brainstorming session to develop strategies to use to address the challenges and a presentation of the resources available to support them.

Between Fall 2017 and Spring 2019, the ER training workshop was implemented four times. After each implementation, iterative refinements were made based on formative evaluation data collected from workshop participants and reflections shared by workshop facilitators. These changes included increased time for practice facilitation, the addition of an equity and inclusion awareness activity as a practice facilitation activity, more time dedicated to the discussion of evaluation and assessment tools, refinement of the preworkshop and day 2 preparation assignments, and incorporation of a structured backward design packet to guide the curricular development work time on day 2.

This study of workshop effectiveness was guided by the following evaluation questions:

EQ2.1: What is the perceived value and quality of the workshop?

EQ2.2: Did the workshop meet its stated goals?

EQ2.3: What were participants' self-reported gains in confidence?

EQ2.4: Do these workshops provide opportunities for participants to make sufficient progress on developing ER curricula and implementation plans for their interventions?

EQ2.5: What is the longer-term impact of the facilitator training workshop on participants' capacity to successfully implement ER interventions?

EQ2.6.: Are trained facilitators implementing their planned ER interventions? In what ways is the ER curriculum being implemented in these interventions?

EQ2.7: What do workshop participants perceive the quality of their intervention implementations to be?

S2 METHODS

Participants

Invitations to participate in the workshop were distributed throughout the NRMN community, to all National Science Foundation-funded Research Experiences for Undergraduate site program directors, and to graduate and undergraduate training program directors of NIH-funded Minority Access to Research Careers and NIH-funded R25 programs. Individuals who applied to attend a workshop and intended to implement ER training with undergraduate and/or graduate trainees within 12 months of attending the workshop were invited to participate. Four 2-day workshops with 114 participants from 64 unique institutions were held between Fall 2017 and Spring 2019. Nine attendees (8%) were from historically black colleges and universities, and five attendees (4%) were from Hispanic-serving institutions. Participant demographics are

TABLE 4. Facilitator training workshop participant demographic information (N = 94)^a

Gender	Percent
Female	72.3
Male	23.4
Not reported	4.3
Race/ethnicity	
White	69.1
Asian	5.3
African American	11.7
Multiple races	5.3
Other	3.2
Not reported	5.4
Hispanic	11.7
Professional role	
Professor	36.2
Scientist/researcher	12.8
Lecturer/instructor	11.7
Dean	4.3
Training program director	34.0
Postdoctoral fellow	1.1
Graduate student	4.3
Other (e.g., other director and coordinator positions)	19.1
Percentage of job dedicated to implementing mentee training	31.22

^aRespondents could select Hispanic in addition to a race category and could select multiple professional roles. As a result, total percentages for the sample may add up to more than 100%.

reported in Table 4, and the career stages and disciplines of their trainees in Table 5.

Procedure

At the conclusion of each workshop, an evaluation survey was distributed electronically to all workshop participants. Of 114

TABLE 5. Facilitator training workshop participants' research training program information (N = 114)^a

Career stage of research trainees	N	Percent
Undergraduate	88	77.2
Postbaccalaureate	14	12.3
Graduate	40	35.1
Other	8	7.1
Discipline of trainees		
Biological sciences	43	37.7
Computer science	4	3.5
Engineering	17	14.9
Environmental research and education	7	6.1
Math and physical sciences	34	29.8
Social and behavioral sciences	14	12.3
Humanities	3	2.6
Medicine	7	6.1
Biomedical	14	12.3
All STEM disciplines	10	8.8
All disciplines	7	6.1
Other	6	5.3

^aCareer stage of trainees and disciplines of trainees are based on information provided on participant workshop applications. Some individual workshop participants hosted multiple programs for trainees that spanned multiple career stages and disciplines, so column totals may add up to more than 100%.

workshop participants, 94 (82%) responded to the survey. Data from these surveys were used to answer EQs 2.1 through 2.5. Facilitators who attended workshops in Fall 2017 and Spring 2018 (N = 52) were invited to complete a second, follow-up implementation survey in Fall 2018. Thirty facilitators completed the survey (58% response rate). Data from this survey were used to answer EQs 2.5 through 2.7. Copies of both the postworkshop and follow-up implementation surveys are available in the Supplemental Material (IRB protocol 2016-0458).

Data Sources and Analyses

Perceived Quality and Value of the Workshop (EQ 2.1). The quality and value of the workshop were assessed from participants' ratings on a five-point scale ranging from 1 (poor) to 5 (excellent) of several aspects of the workshop (e.g., facilitation of activities, the curriculum, participant materials) and their ratings of the value of specific topics addressed in the workshop (e.g., developing learning objectives, assessing trainee learning) on a five-point scale from 1 (not at all valuable) to 5 (extremely valuable). Means and standard errors for each item were calculated. Items were iteratively refined as the workshop structure and topics evolved.

Workshop Met Its Stated Goals (EQ 2.2). Participants were asked: "To what extent do you feel you met each of the following workshop objectives?" Responses could range from 1 (not at all) to 5 (a great deal). One question, "become familiar with NRMN resources & community" was changed to "become familiar with resources & community available to Entering Research facilitators" starting with the Fall 2018 workshop to acknowledge the continuation of the community beyond the funding period of the NRMN grant. These two items were combined and analyzed together. Means and standard errors for the extent to which participants felt that they met each of the workshop objectives were calculated.

Participants' Self-Reported Gains in Confidence (EQ 2.3). Participants were asked to retrospectively assess their confidence with regard to several skills and practices addressed in the workshop, thinking back to before attending the workshop and now, after the workshop. Responses could range from 1 (not at all confident) to 7 (extremely confident). Means and standard errors for each skill were calculated and dependent-samples *t* tests were conducted to determine whether the differences were statistically significant at $p < 0.05$.

Assessment of Progress on ER Curriculum and Implementation Plans (EQ 2.4). We were particularly interested in whether the "work time" incorporated into day 2 of the workshop was valuable and provided participants with a tangible outcome (i.e., a curriculum and implementation plan). To assess this, we asked participants to consider whether the time spent building their curricula during the workshop was valuable; responses to this question could range from 1 (strongly disagree) to 5 (strongly agree). Participants were also asked to report how complete they felt their curricula and implementation plans were at the end of the workshop on a scale ranging from 0% to 100%. Means and standard errors for the question assessing the value of the time spent on curricular planning as well as the extent to which participants' curricula and implementation

plans were complete at the conclusion of the workshop were calculated.

Long-Term Impact of Workshop on Participant Implementations of ER (EQ 2.5). To evaluate the impact of the workshop on participants' implementations, we considered several sources of data. In the postworkshop evaluation survey, we asked participants, "Have your implementation plans changed as a result of attending this workshop?" and "If yes, please explain any changes in your implementation plans." This invited participants to describe what changes they had made or intended to make based on their experiences in the workshop. First, the percentage of individuals who responded "yes" to the question asking whether their implementation plans had changed as a result of attending the workshop was calculated. Next, the open-ended responses to the question asking them to explain any changes they had made in their implementation plans were analyzed. An open-coding approach was used by one researcher (A.R.B.) to identify themes in the responses. Following the process described in study 1 (Creswell, 2009), these themes, along with the original responses, were reviewed by a second researcher (J.L.B.). Any discrepancies in coding were then discussed until agreement could be reached. The total number of responses assigned to each code was tallied to identify the most prominent themes. The final codes, along with definitions and example responses, are reported in the *S2 Results and Discussion*.

On the follow-up implementation survey sent to participants in the first two workshops, participants were asked: "How helpful was attending the Facilitating Entering Research workshop in preparing you to implement this training?" Responses could range from 1 (extremely unhelpful) to 5 (extremely helpful). Also, as part of the follow-up implementation survey, participants were asked how confident they felt in their ability to implement ER training activities, thinking back to before attending the workshop and now, after implementation. Responses could range from 1 (not at all confident) to 7 (extremely confident). Means and standard errors were calculated to determine the average facilitator self-reported level of preparation and confidence to implement a training intervention; a dependent-samples *t* test was conducted to determine whether participants' retrospective confidence ratings were significantly different from the confidence that they reported before attending the workshop.

Implementations of ER by Trained Facilitators (EQ 2.6). We asked ER training workshop participants whether they had actually implemented the ER interventions they planned during the workshop or some version of them on the follow-up implementation survey. The response options were "yes," "no," and "no, but have plans to implement." Then participants were asked to list the number of stand-alone workshops; workshop series; courses/seminars; and summer research programs they had implemented or planned to implement in 2017, 2018, and 2019 using ER activities. To understand how these implementations were supported, we asked whether their implementation was institutionally funded, extramurally funded, or other, with the invitation to provide more information. Frequencies were calculated for participants' responses to questions asking whether or not they had implemented any training interventions using ER curricular materials, the different types of inter-

ventions that they had implemented or planned to implement between 2017 and 2019, and whether their programs were institutionally or extramurally funded.

To begin to track which ER activities and which areas of trainee development in the ER conceptual framework were being incorporated into research trainee interventions, we asked participants to report the percentage of their implementations that consisted of activities from the ER curriculum and how often their trainings addressed each area of trainee development outlined in the ER curriculum. Responses could range from 1 (never) to 5 (always). Means and standard errors for the extent to which ER activities were included in implementations and the extent to which facilitators were addressing each area of trainee development in the ER conceptual framework were calculated.

Quality of ER Implementations (EQ 2.7). Finally, to assess whether participants believed they had the capacity to successfully implement their training interventions, we asked them to rate the overall quality of the implementations they had done thus far from 1 (very low) to 5 (very high). The mean and standard error for responses were calculated.

S2 RESULTS AND DISCUSSION

EQ2.1: Participants Valued the Workshop Activities

Participants provided ratings of several different aspects of the workshop (Supplemental Figure S1) and nearly all participants (95%) rated the overall quality of the workshops as very good or excellent ($n = 91$). Participants also reported that they valued the specific workshop activities (Supplemental Table S1), with the perceived value of activities ranging from a minimum of 3.43 ($SE = 0.130$) to a maximum of 4.52 ($SE = 0.209$) on a scale ranging from 1 (not at all valuable) to 5 (extremely valuable). Based on feedback from workshop participants and reflections from the workshop leaders, we made significant revisions between workshops 3 and 4. A curriculum development packet was introduced and the work time on day 2 was restructured into a series of smaller, more focused tasks. The feedback on these revisions showed that participants valued the added structure (range of mean value: 3.60–3.92). In fact, the overall perceived value of the curricular development work time on day 2 increased from 4.25 ($SE = 0.097$) average for workshops 1 through 3 to 4.52 ($SE = 0.209$) for workshop 4.

Revision of the facilitator training workshop based on evaluation survey results from participants and feedback from the workshop facilitators continues. In particular, we are keenly interested in the feedback from follow-up surveys sent after participants have implemented their interventions. We recognize that it is not until participants actually implement the custom curricula they develop that they can fully appreciate the value of what they learned in the workshop and, more importantly, identify what was missing from the workshop that would have helped them be more successful.

EQ2.2: Participants Reported That the Facilitator Training Workshops Achieved Their Stated Objectives

The evaluation data provide evidence that the facilitator training workshops met the stated goals (Supplemental Figure S2). Participant responses on the postworkshop survey indicated

that the workshop was most effective in helping them to identify areas of trainee development for their programs ($M = 4.29$; $SE = 0.087$) and least effective in helping them select evaluation and assessment tools for their programs ($M = 3.19$; $SE = 0.112$).

We realized after the first two workshops that the backward design process was new to most of the participants, particularly those who had been trained as natural scientists. Most of the participants had never formally articulated learning objectives for their research trainees, but they had overarching program goals and very sophisticated ideas about what it takes to become successful researchers in their fields. The facilitator training workshop served as an opportunity to formalize and prioritize their trainee learning objectives. They reflected on what they were already doing in their programs and courses to support their trainees to achieve the objectives, identified what was missing from their programs, and used the ER activities to fill the gaps. We believe the high evaluation ratings about identifying areas of trainee development for their programs reflect this experience.

The lower evaluation ratings of the evaluation and assessment tool activities on the other hand may be attributed to at least two factors. First, the development of the ERLA and program evaluation tools were occurring in parallel with the workshops, and the ERLA was not validated or publicly available until workshop 3 (Fall 2018). The tools are now published (Branchaw *et al.*, 2020) and available through the online assessment portal at the Center for the Improvement of Mentored Experiences in Research (CIMER). Consequently, we anticipate that satisfaction with this part of the workshop will increase in the future. The second factor to which we attribute the lower evaluation ratings for the assessment and evaluation activities in the workshop is the limited time that participants have to work on their custom curricula. Though most workshop attendees reported having made significant progress in developing their custom curricula at the end of the workshop

(EQ2.4), the selection of evaluation/assessment tools is typically left until the end, and many did not get to it. Based on this, we are considering increasing the length of the workshop or, as suggested by several participants, assigning preworkshop activities that get participants started on curricular development before they arrive.

EQ2.3: Participants Reported Significant Gains in Confidence as a Result of Attending These Workshops

Using retrospective confidence gains scales, we were able to assess the degree to which participants' confidence across five areas changed as a result of attending the workshop (Figure 4). We found that facilitators reported significant gains in all areas, including their ability to use the ER curricular activities and supporting resources ($t(86) = 22.01$, $p < 0.001$); to facilitate mentee training using the process-based approach ($t(86) = 17.178$, $p < 0.001$); to implement research mentee training at their home institution ($t(85) = 15.737$, $p < 0.001$); to use metrics and tools to assess the effectiveness and impact of research mentee training ($t(87) = 12.087$, $p < 0.001$); and to understand the implementation process ($t(85) = 13.342$, $p < 0.001$).

The increases in confidence reported by the workshop participants are consistent with the data showing that the workshop met the stated objectives (EQ2.2), and they support the conclusion that offering the workshop is an effective way to support research training program directors to use the ER resources. Because the second edition of the ER curriculum and assessment and evaluation tools is designed for flexibility and use in any undergraduate or graduate research training program, their effective use requires program directors to have the confidence to adapt the materials in ways that will meet their research trainees' specific needs. These data support that the facilitator training workshop develops this confidence and therefore can be used to train program directors participating in future research.

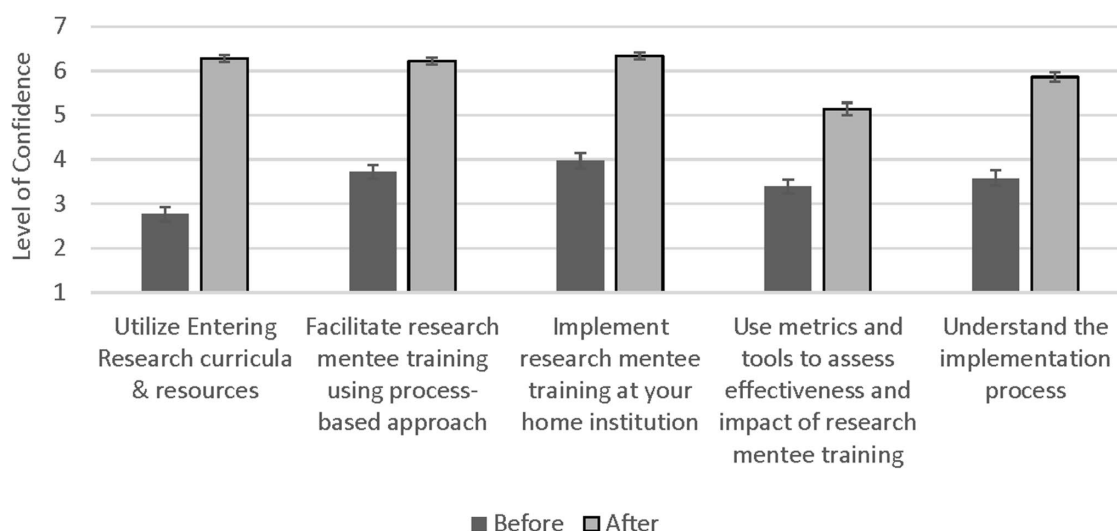


FIGURE 4. ER facilitator training workshop participant self-reported gains in confidence. Workshop participants were asked to retrospectively rate their confidence before and after the workshop. Error bars represent the SEM. Responses ranged from 1 (not at all confident) to 7 (completely confident). All confidence gains were significant ($p < 0.001$). $N = 86$ to 88 .

EQ2.4: Participants Reported Making Progress in Developing Their Custom Curricula and Implementation Plans and They Valued the Dedicated Work Time

Participants felt that the time spent developing their custom curricula and implementation plans was valuable. Eighty-four percent of participants indicated that they agreed or strongly agreed with this statement ($M = 4.35$; $SE = 0.074$). At the end of the workshop, participants reported that the curricula they developed during the workshop were on average 68% complete and their intervention implementation plans were 66% complete. Completeness of curricular plans ranged from 19% to 100%, while completeness of implementation plans ranged from 10% to 100%.

Based on these results, we have concluded that the dedicated work time on day 2 is time well spent. We hypothesize that the variability in reported completeness of the custom curricula and implementation plans was influenced by whether workshop participants were integrating ER activities into an existing program or developing a new program, with those in the latter category making less progress. We began to collect data to test this hypothesis in the third and fourth workshops. Those data show that, on average, workshop participants who were starting a new curricular plan ($n = 20$) report that their curricula and implementation plans were 56% and 54% complete, respectively, at the end of the workshop; those who were working from a draft syllabus or curricular plan ($n = 15$) had curricula and implementation plans that were 67% and 66% complete; and those who were modifying a complete curriculum or plan ($n = 12$) were 79% and 75% complete. We will continue to monitor curriculum and implementation plan completeness and to refine the workshop to maximize progress made by participants at varying starting points.

EQ2.5: Workshop Participation Led to Changes in Implementation Plans and Increased Confidence in Ability to Implement Training

Across all four workshops, the majority of participants (80%) reported that their implementation plans changed as a result of attending the workshop. When asked to describe the changes, 66 participants provided a response. Based on an open-coding approach, nine themes were identified (Table 6). The majority of codes assigned related to the content of the implementation (e.g., incorporation of ER activities or topics into the implementation, $n = 31$; 47% of responses) and the structure of the implementation (e.g., the number of sessions or the structure of the sessions themselves, $n = 28$; 42% of responses).

Participants commented on how the activities helped guide their implementation plans: “I more clearly see [the] gaps in our curriculum and will use ER2 to address them.” Responses like this one demonstrate that the workshop provided an opportunity for participants to revisit their program objectives in comparison to the ER conceptual framework and the activities that align with it to identify areas where their programs could be expanded or changed to better meet the needs of their research trainees.

With regard to structure, many respondents noted that they would change the scope or ordering of their programming based on what they had learned at the workshop. One workshop participant reported the intention to shift the programming from the individual research groups to a more cross-group

approach: “Most of the mentee training was left to individual research groups apart from a couple of professional development workshops that we participated in with partner organizations on campus. Now, I have many activities I can customize for our group and can implement right away at the start of the program rather than being limited by other partner group’s [sic] calendar of events.” Another workshop participant noted “We will expand the scope of our curriculum based on the resources we got from this workshop.”

When asked on the follow-up implementation survey to think back to attending the workshops in Fall 2017 and Spring 2018, all respondents noted that the workshop was either helpful or extremely helpful to them in preparing to implement their training interventions ($M = 4.71$; $SE = 0.074$). Seventy-six percent of respondents reported their level of preparation to implement after attending the workshop as either a lot or a great deal. This high level of preparation was also reflected in retrospective ratings of confidence in their ability to implement their training interventions. On a scale ranging from 1 (not at all confident) to 7 (extremely confident) facilitators reported their levels of confidence as 4.00 ($SE = 0.244$) before the training, and 6.37 ($SE = 0.132$) after the training, $t(26) = 9.66$, $p < 0.001$.

EQ 2.6: Trained Facilitators Are Implementing ER Activities in Their Training Interventions

On the follow-up implementation survey, 26 facilitators (88%) reported that they had implemented some form of training using ER activities since attending the workshop. Of the remaining four facilitators, three had plans to implement, and one did not respond to the question. Figure 5 summarizes the number of implementations completed and planned from 2017 to 2019. On average, respondents had implemented ER activities in their training interventions at least once in 2017 and 2018 and had another implementation planned for 2019. Sixty-seven percent of facilitators reported that their implementations were institutionally funded or part of a class offered by their institution; 47% reported that their implementations were extramurally funded; and two facilitators (7%) reported that their implementations were not funded.

Workshop participants who implemented their training interventions and responded to our survey indicated that, on average, 67% of the activities they used in their implementations were from the ER curriculum. We also asked them to report how often their training interventions addressed each of the areas of trainee development in the ER conceptual framework. Responses could range from 1 (never) to 5 (always) (Figure 6). The percentage of responses of most of the time (4) or always (5) were highest for the areas of Research Comprehension and Communication Skills (70%) and Professional and Career Development Skills (53%), with lower percentages reported for Researcher Confidence and Independence (47%), Practical Research Skills (36%), Equity and Inclusion Awareness and Skills (33%), Researcher Identity (33%), and Research Ethics (17%).

Though a high percentage of survey respondents reported that they had implemented some form of ER training since attending the workshop (88%), the response rate to the follow-up survey was only 58%, so we do not know whether the remaining workshop alumni have or plan to implement the curricula they developed. However, the information provided

TABLE 6. Codes and example responses of changes to implementation plans

Code and definition	Example responses	Number of times code assigned ^a
Activities. Mentions incorporation of ER activities into curricular plan.	<ul style="list-style-type: none"> • “There are several activities that I will implement related to posters, mentor relationship, issues in a research setting.” • “We added a session on the importance of diversity [in] STEM research and learning.” 	31
Structure. Changes in number, timing, organization, or execution of sessions.	<ul style="list-style-type: none"> • “I had previously planned to rely heavily on panels, rather than facilitate activities, and now I have switched to 80% <i>Entering Research</i> activities.” • “Decided to make it a year long course instead of a 10-week long course.” 	28
Progress. Increased awareness of importance of implementation plan, or describes progress on an implementation plan.	<ul style="list-style-type: none"> • “I did not know what an implementation plan was. I learned a lot about the entire process of facilitating. I feel since there wasn’t any initial plan ... establishing one is definitely a change.” • “Came with only a vague idea for a program, leaving with a concrete plan.” 	16
Assessment/Evaluation. Changes in evaluation or assessment plans.	<ul style="list-style-type: none"> • “I changed the ordering of my activities and added additional opportunities for assessment.” • “I also hope to update the evaluations to focus on gains rather than ‘pre and post measures’ and some of the [activities].” 	6
Learning Objectives. Mention of change in learning objectives or consideration of learning objectives in implementation planning.	<ul style="list-style-type: none"> • “I was able to refocus my learning objectives and adjust the curriculum accordingly.” • My prior implementation plans were more ad hoc. I now plan to be more intentional in learning objectives and assessment, which will feed into the implementation plan.” 	4
Facilitation. Incorporating facilitation as a technique into the implementation plan.	<ul style="list-style-type: none"> • “I will do more facilitation and less lecture.” • “The curriculum I inherited has some facilitation, but I will be adding more.” 	4
Partners. Mention of partnering with other departments or individuals on campus.	<ul style="list-style-type: none"> • “I realized that I needed to bring in additional campus partners. I am now looking for ... existing resources.” • “I will involve more faculty.” 	3
Recruitment/Participation. Mention of recruitment or targeting specific participants.	<ul style="list-style-type: none"> • “I am considering additional ways to increase participation.” 	1
Audience. Change in target audience.	<ul style="list-style-type: none"> • “It has changed who I will focus on meaning will [implement] summer and academic year.” 	1

^aN = 66. Codes are presented in order of most frequently to least frequently assigned. Responses could be assigned multiple codes, so total number of codes may exceed the total number of responses.

by the individuals who did respond to the survey suggests that the ER activities can be implemented in institutional and grant-funded programs and that facilitators are successfully embedding these activities into implementations that incorporate other activities and topics. The results also provide information about which activities are being implemented and therefore potentially contributing to various trainee outcomes that will be tested in future research.

EQ 2.7: Facilitators Believe That They Are Successfully Implementing ER Activities in Their Training Interventions

The majority of the workshop participants (73%) rated the quality of their postworkshop implementations as either high or very high, which aligns with their reported high levels of confidence in their ability to implement. These data provide evidence that the workshop was effective in preparing research training program directors to develop and implement a training intervention. However, this is all self-assessment data, so implementation evaluation and learning outcomes data from the research trainees in their programs will be collected in the future to more objectively measure the quality of the implementations.

S2 CONCLUSIONS

Based on the evaluation data collected, we conclude that the ER facilitator training workshop successfully addressed the workshop goals and provided participants with confidence in their ability to design a curriculum using ER activities, identify assessment and evaluation tools to measure the impact of their implementations of those activities, and plan their training interventions. Though we continue to refine this workshop, the evaluation data reported here give us confidence that we can use it to train facilitators to implement the ER activities with the level of fidelity needed to conduct in-depth investigations of the impact of the ER curriculum on research trainee learning and development in future research studies.

SUMMARY

Together, the data from these two design and development research studies provide evidence that the second edition of the ER curriculum is ready for use by undergraduate and graduate research training program directors; that it is sufficiently promising to warrant more systematic, in-depth efficacy, effectiveness, and scale-up research in the future; and

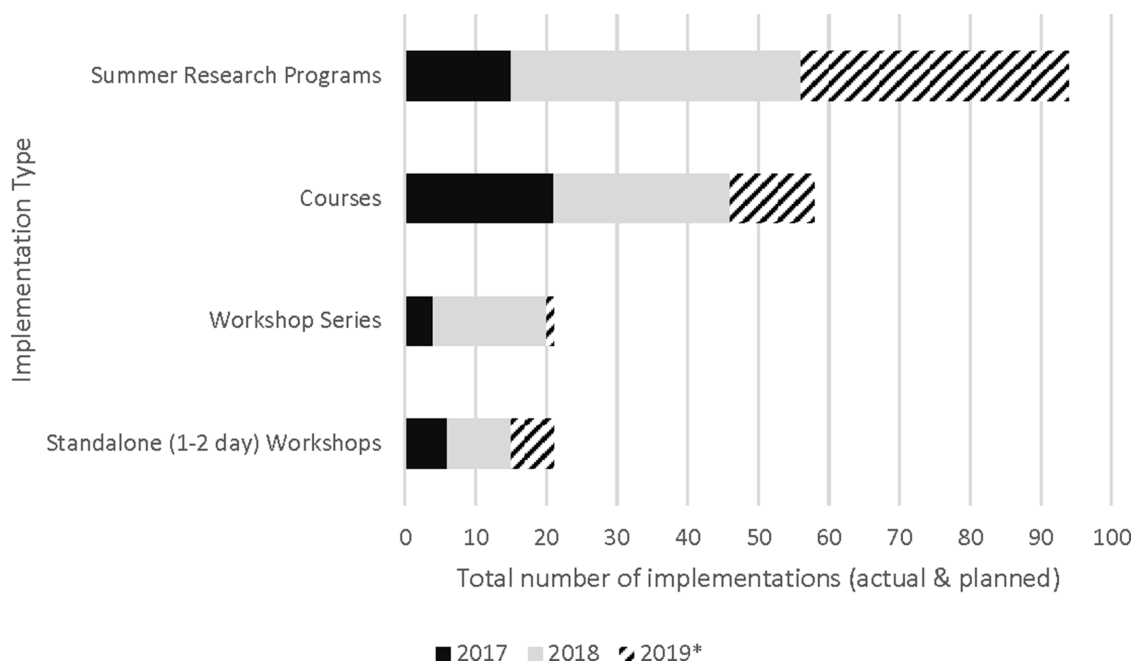


FIGURE 5. Previous and planned implementations for trained facilitators of ER. Facilitators were asked to report the number of trainings they had implemented in each category during 2017 and 2018 and had planned for 2019. *2019 trainings had not yet occurred at the time this question was asked.

that the ER training workshop is an effective way to train research program directors to use the curriculum in these future studies. Study 1 provides evidence that the curriculum can be successfully implemented in a wide variety of formats with different combinations of activities. Facilitators found the implementation guides and trainee materials useful in their

implementations, and facilitators and trainees rated the activities favorably. Study 2 provides evidence that the ER facilitator training workshop increases participants' knowledge, facilitation skills, and confidence in their ability to implement the custom curricula they create in various types of implementations and with trainees at various career stages.

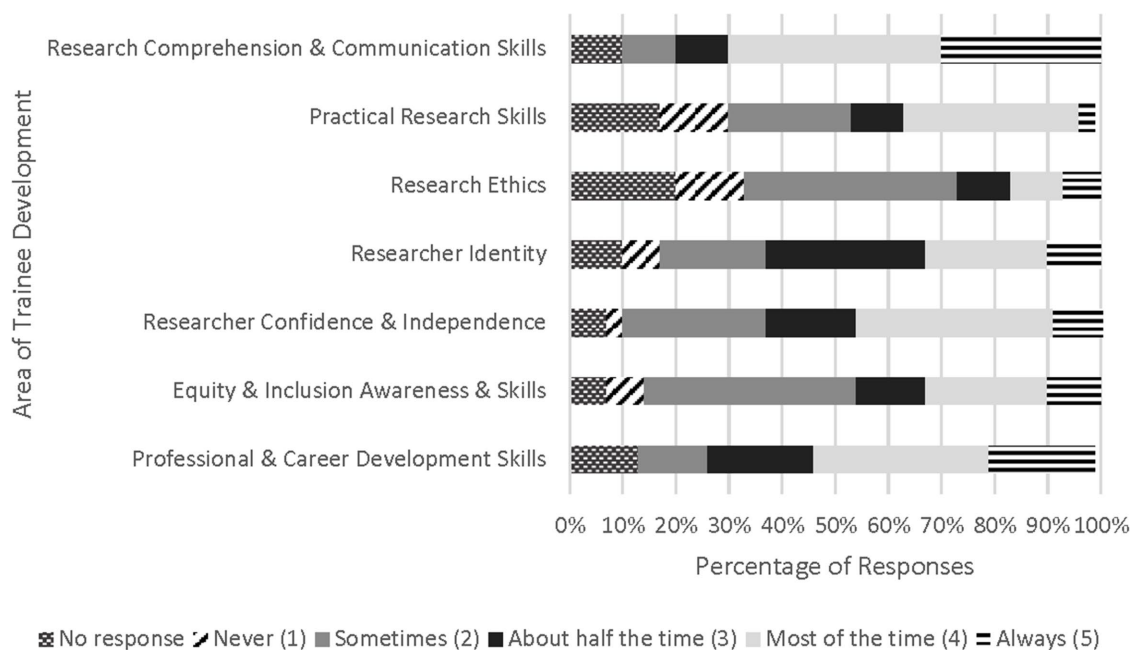


FIGURE 6. Areas of trainee development addressed in trained facilitator implementations. The frequency with which each area of trainee development was addressed in trainings that were implemented by ER facilitator training workshop participants. $N = 24-28$.

LIMITATIONS AND FUTURE RESEARCH

The limitations of our design and development research studies present several opportunities for future research. Study 1 was designed to determine whether research training program directors would be willing and able to facilitate the second edition ER curricular activities and whether research trainees would perceive participating in the curricular activities as valuable. Self-reported data provided by trainees and facilitators were the only source used to assess this. Additional measures of trainee outcomes (such as research posters assessed using rubrics) would have provided more robust evidence of the impact of the curriculum on trainee learning and skill gains and would have guarded against potential socially desirable responses. Rubrics developed to assess such products are available in the curriculum and will be used to more objectively assess trainee outcomes in future research.

The present study was not designed to test causal inference; we did not measure actual behavior changes or the long-term impact of the curriculum on trainee development. Research measuring these variables is needed to determine whether the positive self-reported trainee outcomes lead to actual trainee development and persistence in research. However, the evidence reported in study 1 does show that the curriculum can be implemented with fidelity and that research trainees find participation valuable, which lays the foundation for future efficacy, effectiveness, and scale-up research studies investigating the long-term and potentially causal impacts of the curriculum on trainee development and persistence.

Another limitation of study 1 was that all of the ER curricular activities were not tested equally, and a few activities were not tested at all. This was a function of the design and intent behind study 1. We wanted to evaluate whether research training program directors would be willing to use the curriculum, and we wanted their use to reflect how program directors would actually engage with the curriculum, so we allowed them to select their own activities, rather than assign specific activities to each pilot testing site. This resulted in extensive testing of some activities and minimal testing of others, but provided insight into which areas of trainee development research training program directors are most interested in addressing and which activities they are comfortable facilitating. The uniform template and rigorous review process used to develop the activities gave us confidence that the positive evaluations reported about the structure and content of the tested activities will apply to the activities that have not yet been tested. However, additional research is needed to confirm this. Future research studies will test activities in specific areas of trainee development implemented in specific contexts to allow us to gather evidence to confirm these assumptions and to test for causal relationships and for relationships between individual activities and attributions of learning. In addition, more controlled studies will allow us to administer fewer surveys in semester-long and yearlong implementations, which should yield higher and more consistent response rates. When comparison groups are used, we will also be able to compare the impact of the ER curriculum on trainee outcomes to the outcomes from structured training programs that do not implement the ER curriculum.

Study 2 was designed to measure the effectiveness of the ER training workshop to determine whether it could be used in

future, in-depth research studies to train research training program directors participating in the studies. A limitation of study 2 was that data were collected as the workshop content and process were in development, so specific workshop components changed from one implementation to the next. Also, the follow-up survey was only sent to participants in the first two workshops, and some of the changes, most notably the incorporation of a curricular development packet to structure and guide participants' work on day 2 and access to the ERLA, were not part of the training workshop they attended. Though these changes were documented and accounted for in the data analysis, there is no way to know how they impacted the overall workshop ratings. Additional follow-up data from participants are needed to confirm the present findings.

Another limitation of study 2 is that evaluation data were not gathered directly from student participants in the postworkshop implementations; thus we could not verify the implementation quality ratings reported by facilitators. Though this was beyond the scope of work funded during development of the workshop, future research studies will include collection of implementation evaluation data from student participants.

In summary, the results of the design and development studies reported here will inform future efficacy, effectiveness, and scale-up research investigating the long-term impacts of the second edition ER activities on research trainee learning, development, and career outcomes. Future research will investigate the impact that different ER activities or sets of activities have on different populations of students at different career stages and types of institutions. Furthermore, these studies will test the hypothesis that using the second edition ER curriculum has a positive impact on recruiting and retaining more diverse populations of trainees to undergraduate and graduate research training programs.

Ongoing Curricular Development

Development of the ER curriculum continues through the Wisconsin Institute for Science Education and Community Engagement (WISCIENCE; <https://wiscience.wisc.edu/program/entering-research>). In particular, adaptations for use and pilot testing of ER activities in CUREs is planned. The authors invite submission of new activities in any area of trainee development for review, pilot testing, and ultimately publication on the Center for the Improvement of Mentored Experiences in Research training materials website (CIMER; www.cimerproject.org/#). Individuals interested in contributing activities for pilot testing may contact the authors.

Access to the Curriculum

Research training program directors interested in using the second edition of the ER curriculum can access the activities and some premade curricula for common implementations for free on the CIMER website or by purchasing the *Entering Research* book (Branchaw *et al.*, 2020) from Macmillan Publishing (<https://store.macmillanlearning.com/us>). Each activity contains trainee materials and detailed facilitator instructions to guide implementation. Research training program directors interested in designing a custom curriculum are invited to participate in an ER facilitator training workshop (<https://cimerproject.org/training-fer>) and can find instructions about using the backward design process in the book.

ACKNOWLEDGMENTS

Funding for this research was provided by the National Institutes of Health (NIH U54MD009479/ NIGMS U54GM119023) and the Department of Kinesiology and WISCIENCE at the University of Wisconsin–Madison. The work is solely the responsibility of the authors and does not represent the official view of the NIH or the University of Wisconsin–Madison. A special thanks to the trainees and mentors who participated in this NIH Diversity Program Consortium study and to our friendly reviewers, Molly Carnes, Elizabeth Meyerand, and Jennifer Gleason, who provided valuable feedback on a draft of this article.

REFERENCES

- Anderson, D. D., & Shore, W. J. (2008). Ethical issues and concerns associated with mentoring undergraduate students. *Ethics & Behavior*, 18(1), 1–25.
- Balster, N., Pfund, C., Rediske, R., & Branchaw, J. (2010). *Entering Research: A course that creates community and structure for beginning undergraduate researchers in the STEM disciplines*. *CBE—Life Sciences Education*, 9, 108–118.
- Bartlett, J. (2012). *A model role evaluation of mosaic mentoring programmes*. London: Demos.
- Bauer, K. W., & Bennett, J. S. (2003). Alumni perceptions used to assess undergraduate research experience. *Journal of Higher Education*, 74(2), 210–230.
- Branchaw, J. L., & Butz, A. R. (2019, April). The Entering Research Learning Assessment (ERLA): Measuring research trainee learning and development. Poster presented at: 11th Conference on Understanding Interventions that Broaden Participation in Science Careers (Baltimore, MD).
- Branchaw, J. L., Butz, A. R., & Smith, A. R. (2020). *Entering Research: A curriculum to support undergraduate and graduate research trainees* (2nd ed). New York: Macmillan.
- Branchaw, J. L., Pfund, C., & Rediske, R. (2010). *Entering research: Workshops for students beginning research in science*. New York: Freeman.
- Butz, A. R., & Branchaw, J. L. (2020). Entering Research Learning Assessment (ERLA): Validity evidence for an instrument to measure undergraduate and graduate research trainee development. *CBE—Life Sciences Education*, 19(2) (in press).
- Butz, A. R., Spencer, K., Thayer-Hart, N., Cabrera, I. E., & Byars-Winston, A. M. (2018). Mentors' motivation to address race/ethnicity in research mentoring relationships. *Journal of Diversity in Higher Education*, 12(3), 242–254. <https://doi.org/10.1037/dhe0000096>
- Byars-Winston, M., Gutierrez, B., Topp, S., & Carnes, M. (2011). Integrating theory and practice to increase scientific workforce diversity: A framework for career development in graduate research training. *CBE—Life Sciences Education*, 10(4), 357–367.
- Carter, F. D., Mandell, M., & Maton, K. I. (2009). The influence of on-campus, academic year undergraduate research on STEM Ph.D. outcomes: Evidence from the Meyerhoff Scholarship Program. *Educational Evaluation and Policy Analysis*, 31(4), 441–462.
- Carver, S., Van Sickle, J., Holcomb, J., Quinn, C., Jackson, D., Resnick, A., ... & Marquard, A. (2017). Operation STEM: Increasing success and improving retention among first-generation and underrepresented minority students in STEM. *Journal of STEM Education*, 18(3), 30–39.
- Chan, A. W. (2008). Mentoring ethnic minority, pre-doctoral students: An analysis of key mentor practices. *Mentoring & Tutoring: Partnership in Learning*, 16(3), 263–277.
- Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching*, 51(5), 555–580.
- Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *Journal of Social Issues*, 67(3), 469–491.
- Chesler, M. A., & Lohman, J. E. (1971). Changing schools through student advocacy. In Schmuck, R., & Miles, M. (Eds.), *Organization development in schools* (pp. 185–211). Palo Alto, CA: National Press Books.
- Crane, C., McKay, T., Mazzeo, A., Morris, J., Prigodich, C., & de Groot, R. (2011). Cross-discipline perceptions of the undergraduate research experience. *Journal of Higher Education*, 82(1), 92–113.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed-methods approaches*. Thousand Oaks, CA: Sage.
- Darling, L. W. (1986). The mentoring mosaic: A new theory of mentoring. In Gray, W. A., & Gray, M. M. (Eds.), *Mentoring: Aid to excellence in career development, business, and the professions* (Vol. 2, pp. 1–7). Vancouver, BC: International Association for Mentoring.
- Davidson, M. N., & Foster-Johnson, L. (2001). Mentoring in the preparation of graduate researchers of color. *Review of Educational Research*, 71, 549–574. <https://doi.org/10.3102/00346543071004549>
- Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal*, 50, 683–713. doi: 10.3102/0002831213482038
- Eby, L. T., Butts, M., Lockwood, A., & Simon, S. A. (2004). Protégés' negative mentoring experiences: Construct development and nomological validation. *Personnel Psychology*, 57(2), 411–447.
- Fenning, K. (2004). Cohort based learning: Application to learning organizations and student academic success. *College Quarterly*, 7(1). Retrieved February 1, 2019, from <http://collegequarterly.ca/2004-vol07-num01-winter/fenning.html>
- Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2016). Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice. *Journal of Engineering Education*, 105(2), 312–340.
- Hathaway, R. S., Nagda, B. A., & Gregerman, S. R. (2002). The relationship of undergraduate research participation to graduate and professional education pursuit: An empirical study. *Journal of College Student Development*, 43(5), 1–18.
- Head, F. A., Reiman, A. J., & Thies-Sprinthall, L. (1992). The reality of mentoring: Complexity in its process and function. In Bey, T. M., & Holmes, C. T. (Eds.), *Mentoring: Contemporary principles and issues* (pp. 5–24). Reston, VA: Association of Teachers.
- Hunter, A. B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in cognitive, personal and professional development. *Science Education*, 91(1), 36–74.
- Hurtado, S., Eagan, M. K., Cabrera, N. L., Lin, M. H., Park, J., & Lopez, M. (2008). Training future scientists: Predicting first-year minority student participation in health science research. *Research in Higher Education*, 49(2), 126–152.
- Hurtado, S., Milem, J. F., Clayton-Pedersen, A. R., & Allen, W. R. (1998). Enhancing campus climates for racial/ethnic diversity: Educational policy and practice. *Review of Higher Education*, 21, 279–302.
- Institute of Education Sciences and the National Science Foundation. (2013). *Common guidelines for education research and development: A report from the Institute of Education Sciences, U.S. Department of Education, and the National Science Foundation*. Retrieved February 1, 2019, from www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf
- Junge, B., Quiñones, C., Kakietek, J., Teodorescu, D., & Marsteller, P. (2010). Promoting undergraduate interest, preparedness, and professional pursuit in the sciences: An outcomes evaluation of the SURE program at Emory University. *CBE—Life Sciences Education*, 9(2), 119–132.
- Kim, S. H., Smith, R. H., & Brigham, N. L. (1998). Effects of power imbalance and the presence of third parties on reactions to harm: Upward and downward revenge. *Personality and Social Psychology Bulletin*, 24(4), 353–361.
- Laursen, S., Hunter, A.-B., Seymour, E., Thiry, H., & Melton, G. (2010). *Undergraduate research in the sciences: Engaging students in real science*. San Francisco, CA: Jossey-Bass.
- Layton, R. L., Brandt, P. D., Freeman, A. M., Harrell, J. R., Hall, J. D., & Sinche, M. (2016). Diversity exiting the academy: Influential factors for the career choice of well-represented and underrepresented minority scientists. *CBE—Life Sciences Education*, 15(3), ar41. doi: 10.1187/cbe.16-01-0066
- Lewis, V., Martina, C. A., McDermott, M. P., Trief, P. M., Goodman, S. R., Morse, G. D., ... & Ryan, R. M. (2016). A randomized controlled trial of mentoring interventions for underrepresented minorities. *Academic Medicine*, 91(7), 994–1001.

- Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, 347, 1261757. Retrieved February 1, 2019, from <http://science.sciencemag.org/content/347/6222/1261757>
- Lopatto, D. (2007). Undergraduate research experiences support science career decisions and active learning. *CBE—Life Sciences Education*, 6(4), 297–306.
- Mau, W.-C. J. (2016). Characteristics of U.S. students that pursued a STEM major and factors that predicted their persistence in degree completion. *Universal Journal of Educational Research*, 4(6), 1495–1500.
- Mullen, C. A. (2007). Naturally occurring student-faculty mentoring relationships: A literature review. In Allen, T. A., & Eby, L. T. (Eds.), *Blackwell handbook of mentoring: A multiple perspectives approach* (pp. 119–138). Oxford: Blackwell Publishing.
- Museum, S. D., & Liverman, D. (2010). High-performing institutions and their implications for studying underrepresented minority students in STEM. *New Directions for Institutional Research*, 148, 17–27.
- National Academies of Sciences, Engineering, and Medicine (NASEM). (2015). *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. Washington, DC: National Academies Press. <https://doi.org/10.17226/21851>.
- NASEM. (2017). *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. Washington, DC: National Academies Press. <https://doi.org/10.17226/24622>.
- NASEM. (2018). *Graduate STEM education for the 21st century*. Washington, DC: National Academies Press. <https://doi.org/10.17226/25038>.
- Ngassa, F. N. (2013). Mentoring undergraduate research: Opportunities and challenges. In *Developing and maintaining a successful undergraduate research program* (ACS symposium series 1156, pp. 39–50). Washington, DC: American Chemical Society.
- Packard, B. W.-L. (2015). *Successful STEM mentoring initiatives for underrepresented students: A research-based guide for faculty and administrators*. Sterling, VA: Stylus.
- Prunuske, A. J., Wilson, J., Walls, M., & Clarke, B. (2013). Experiences of mentors training underrepresented undergraduates in the research laboratory. *CBE—Life Sciences Education*, 12, 403–409. doi: 10.1187/cbe.13-02-0043
- Research Triangle International (RTI). (2018). *Focus groups conducted for the Graduate NRC report*. Retrieved January 27, 2019, from http://sites.nationalacademies.org/cs/groups/pgasite/documents/webpage/pga_186164.pdf
- Sabitini, D. A. (1997). Teaching and research synergism: The undergraduate research experience. *Journal of Professional Issues in Engineering Education and Practice*, 123(3), 98–102.
- Seymour, E., Hunter, A. B., Laursen, S. L., & DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education*, 88, 493–534.
- Thiry, H., Laursen, S. L., & Hunter, A. B. (2011). What experiences help students become scientists? A comparative study of research and other sources of personal and professional gains for STEM undergraduates. *Journal of Higher Education*, 82(4), 358–389.
- Wiggins, G. P., McTighe, J., Kiernan, L. J., & Frost, F. & Association for Supervision and Curriculum Development.(1998) *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Wisker, G., Robinson, G., & Shacham, M. (2007). Postgraduate research success: Communities of practice involving cohorts, guardian supervisors and online communities. *Innovations in Education and Teaching International*, 44(3), 301–320.