Contributions Made by Undergraduates to Research Projects: Using the CREDIT Taxonomy to Assess Undergraduate Research Experiences

Matt Honoré,
Oregon Health and Science University

Thomas E. Keller,
Portland State University

Jen Lindwall,
Portland State University

Rachel Crist,
Oregon Health and Science University

Leslie Bienen,
Portland State University

Adrienne Zell
Oregon Health and Science University

Abstract

The authors developed a novel tool, the CREDIT URE, to define and measure roles performed by undergraduate students working in research placements. Derived from an open-source taxonomy for determining authorship credit, the CREDIT URE defines 14 possible roles, allowing students and their research mentors to rate the degree to which students participate in each role. The tool was administered longitudinally across three cohorts of undergraduate student-mentor pairs involved in a biomedical research training program for students from diverse backgrounds. Students engaged most frequently in roles involving data curation, investigation, and writing. Less frequently, students engaged in roles related to software development, supervision, and funding acquisition. Students' roles changed over time as they gained experience. Agreement between students and mentors about responsibility for roles was high.

Keywords

biomedicine; evaluation; mentorship; STEM; undergraduate research

Undergraduate research experiences (UREs) are central to many training programs designed to prepare students for graduate school and future careers in research (Dyer-Barr 2014; Tsui 2007). Previous research has identified a number of potential benefits to students

Adrienne Zell, Oregon Health and Science University, zell@ohsu.edu.

participating in UREs. These benefits include gains in science interest, understanding, skills, confidence, persistence, and career preparation and higher rates of enrollment in postgraduate education (Lopatto 2007; Russell, Hancock, and McCullough 2007; Seymour et al. 2004). Studies provide evidence that UREs can have additional positive effects such as enhancing student retention and encouraging pursuit of graduate education in science, technology, engineering, and mathematics (Eagan et al. 2013; Gregerman et al. 1998). Furthermore, the literature indicates that UREs can be particularly effective in improving educational outcomes for students traditionally underrepresented in STEM fields (Chang et al. 2014; Hurtado et al. 2009). The National Institutes of Health (NIH) defines underrepresented individuals as those from a racial or ethnic background traditionally underrepresented in the health-related sciences (Blacks or African Americans, Hispanics or Latinos, American Indians and Alaska Natives, and Native Hawaiians and other Pacific Islanders); individuals with disabilities, defined as those with a physical or mental impairment that substantially limits one or more major life activities; and individuals from disadvantaged backgrounds. The latter designation is defined as those who meet two or more of the following criteria: were or are homeless; were or are in the foster care system; were eligible for the Federal Free and Reduced Lunch Program for two or more years; did not have parents or legal guardians who completed a bachelor's degree; were or currently are eligible for Federal Pell Grants; received support from the Special Supplemental Nutrition Program for Women, Infants and Children (WIC) as a parent or child; or grew up in a rural area (NIH 2019).

Students with low socioeconomic status (SES) backgrounds have been shown to obtain bachelor's and advanced degrees at significantly lower rates than students from middle and high SES groups (NCES 2015). Past studies have suggested that program effectiveness may be associated with the duration and intensity of UREs, with students developing a more sophisticated understanding of the research process when they have longer research experiences (Thiry et al. 2012).

Given the prevalence of undergraduate research programs and the varied disciplines and institutions that offer them, it is not surprising that these programs offer a wide variety of student activities and experiences. Yet research characterizing these activities and their contributions to student development is limited (Linn et al. 2015). In general, UREs provide exposure to research environments, enabling students to have an apprentice-like experience with hands-on learning and application of research skills. UREs have additional self-efficacy benefits, as students build self-confidence in research settings (Feldman, Divoll, and Rogan-Klyve 2013). Students engaging in UREs typically receive research mentoring to assist with professional socialization and intellectual and skills development and to provide personal support for overcoming challenges and developing confidence (Thiry, Laursen, and Hunter 2011). UREs also may include formal lessons on research ethics and other relevant topics. Ideally, students in UREs participate in multiple phases of the research process to promote an integrated understanding of science (Linn et al. 2015).

Studies investigating UREs often employ measures such as the Survey of Undergraduate Research Experiences (SURE; Lopatto 2007) or the Undergraduate Research Student Self-Assessment (URSSA; Weston and Laursen 2015). These self-report instruments focus on

student perceptions of how much they are learning through their research experiences. This article proposes that another way of understanding and assessing student engagement in UREs is to consider the contributions made by students through their participation in conducting research. This approach derives from theories of experiential, situated learning that emphasize the importance of engaging in authentic tasks within a legitimate community of practice (Thiry, Laursen, and Hunter 2011). In this context, research teams can operate as learning communities in which new members are empowered with agency and responsibility, "which enables everyone to make contributions, even undergraduate students" (Feldman, Divoll, and Rogan-Klyve 2013). This perspective also acknowledges that the work of students in apprenticeship roles can add value to faculty research efforts and enhance scientific productivity (Shortlidge, Bangera, and Brownell 2015).

To assess the contributions made by students during UREs, the authors adapted the Contributor Roles Taxonomy (CRediT), a tool that provides high-level classification of the diverse roles performed in work leading to published research output in the sciences. As an open-source taxonomy, the CRediT tool was designed to provide transparency regarding contributions to scholarly published work and to improve systems of attribution, credit, and accountability (Brand et al. 2015). The CRediT taxonomy consists of 14 roles, including conceptualization, analysis, writing, and funding acquisition (see Table 1). It is intended to enable documentation of contributions beyond generalizations such as "significant" and "legitimate" (Smith and Williams-Jones 2012). Although the CRediT tool was originally conceived and designed with authorship roles in mind, it was hypothesized that this taxonomy could be extended and applied to measure the roles of students in UREs. Two parallel instruments were developed: one survey for students participating in UREs, and the other for their lab mentors or direct supervisors to complete. Each tool listed the 14 taxonomy components as question stems and gave a rating-scale response option ranging from 0 (no responsibility) to 3 (primary responsibility). These measures were referred to as the CREDIT URE (undergraduate research experience).

Assessing responsibilities of students in UREs shifts the focus to specific activities that can be verified by objective observations from others on the research team. This approach also mirrors the way the scholarship and productivity of researchers typically are evaluated. The CREDIT URE, which addresses multiple facets of the research enterprise, can assess not only level of responsibility but also changes in the range of responsibilities over time. Another advantage of focusing on specific research activities is that one can potentially investigate and assess learning processes in UREs by identifying activities that promote certain types of learning and development. Finally, because the CREDIT URE focuses on observable activities and permits parallel responses by student and mentor, it is less susceptible to reference bias (Mathews and Bradle 1983) and social desirability bias (Fisher and Katz 2000) than are self-reported assessments of perceptions of research gains.

This article reports on the use of the CREDIT URE for research and evaluation in the context of an NIH-funded research training program for undergraduates from backgrounds traditionally underrepresented in the biomedical sciences. A core feature of the training is a long-term research placement to work on faculty-directed research projects. To demonstrate

the utility of assessing student contributions in these research settings using the CREDIT URE, the following research questions were addressed:

- In which areas do students make contributions to their UREs?

- To what extent do students have responsibility for contributions in those different areas?

- How do contributions reported by students compare to contributions reported by their mentors?

- Are there observable changes in the nature or degree of contributions over time?

Method

Description

BUILD EXITO is one of 10 demonstration projects in the NIH-funded Building Infrastructure Leading to Diversity (BUILD) initiative to develop and test new approaches for diversifying the future biomedical workforce (Valantine and Collins 2015). The Enhancing Cross-Institutional Training in Oregon (EXITO) initiative is a large, multi-institutional collaboration that provides comprehensive support and training for undergraduates from traditionally underrepresented student populations who aspire to health-related research careers (Richardson et al. 2017). These students receive funding through NIH training mechanisms and are thereafter referred to as student trainees. Key outcomes for EXITO student trainees include persistence in preparation for a research career, graduation, matriculation to and completion of graduate school, and eventual entrance into a biomedical research career. Other psychosocial and research-related outcomes include developing a science identity, acquiring research skills, producing papers and presentations, and seeking grant funding.

The program is referred to as BUILD EXITO to acknowledge that it is part of the BUILD program and maintain its identity as a unique program with a design different than other BUILD programs. BUILD EXITO offers a comprehensive, developmentally sequenced training program featuring an integrated curriculum, research experiences, multifaceted mentoring, and a supportive environment (Keller et al. 2017). Long-term placement in a research learning community (RLC) is a core component of the model. BUILD EXITO RLCs are established based on the willingness and capacity of the research lead (principal investigator) to engage undergraduate students in meaningful research activities on faculty-directed projects. Placements begin in the summer prior to the student's third year and continue for 18 months. BUILD EXITO student trainees are expected to work 10 hours per week during the academic year in the RLCs as compensated trainees, and they have an intensive research experience in their RLCs during the summer between their third and fourth years (30 hours per week). RLC labs are paid by the grant to cover marginal expenses associated with supporting an additional research team member (e.g., staff time, equipment, supplies). BUILD EXITO student trainees are encouraged to remain in one placement throughout the RLC phase of their training with the idea that they will become increasingly integrated into their RLCs; have opportunities to participate in multiple facets of the research process (i.e., study design, data collection, analysis, and reporting); and take ownership

of their specific contributions. RLCs also provide student trainees with opportunities to participate in preparing proposals, writing manuscripts, and giving presentations and to join in other activities as they gain experience and can make greater contributions to the research.

Setting

BUILD EXITO is a collaborative multi-institutional project led by Portland State University, a major public urban university that prioritizes student access and opportunity, and Oregon Health & Science University, a research-intensive academic health center. The BUILD EXITO network includes nine additional partners, a mix of two-year and four-year institutions of higher education spanning Alaska, American Samoa, Guam, Hawaii, the Northern Mariana Islands, Oregon, and Washington. Students entering BUILD EXITO from community college partners eventually transfer to Portland State University to continue the program, whereas students at four-year-university partners complete the entire program at their home institutions. BUILD EXITO institutions hosting RLCs at the time of this study included Oregon Health and Science University (OHSU, research partner); Portland State University (PSU, primary institution); and the four-year EXITO partner universities: University of Alaska–Anchorage (UAA), University of Guam (UG), and University of Hawai'i at M noa (UHM). EXITO RLC placements are organized according to four broad domains representing a range of biomedical disciplines: biological sciences, clinical sciences, community health and social sciences, and chemistry/physics/engineering/ environmental science. These categories are aggregated from a list of majors targeted by the BUILD program.

Sample

BUILD EXITO admitted its first cohort of student trainees in 2015, subsequently adding cohorts in 2016, 2017, and 2018 ($n = 361$). Student trainees in the first three BUILD EXITO cohorts were asked to complete the CREDIT URE as part of this study ($n = 265$). Respondents reflected the diverse populations targeted by the initiative: 67 percent were female, 60 percent were first-generation undergraduate students, 53 percent came from a disadvantaged background, 71 percent received need-based financial aid, and 38 percent were underrepresented minorities. At the time of the most recent CREDIT URE administration, 132 student trainees had been placed in 83 unique RLCs. These students reported spending, on average, 253 hours in their labs over the course of an academic year. In addition to lab hours, student trainees reported spending, on average, 63 hours meeting directly with their research mentors.

Instrument Testing

In the process of developing and refining the CREDIT URE instrument, a draft of the instrument text was sent to seven faculty members for review and comment. Reviewers were OHSU and PSU lab supervisors, unconnected to BUILD EXITO, who represented a variety of disciplines and had undergraduate students working in their labs. The draft included definitions for each level of the taxonomy, and reviewers were asked to make any edits to the definitions that would make them more comprehensible to undergraduates. Faculty research mentors provided high ratings for both face and content validity. Minor revisions were made based on this feedback, including shifting from a binary response option (did/did

not have responsibility) to a scaled response (0 = no responsibility; 1 = little responsibility; 2 = moderate responsibility; 3 = primary responsibility).

Instrument Administration

The CREDIT URE assessment was piloted in spring 2016, after the first cohort of BUILD EXITO RLC student trainees had been in their placements for approximately six months. The survey was emailed to 16 student-mentor pairs via a secure method of survey administration (see Table 2). Three mentors had more than one mentee, resulting in 16 students and 13 faculty receiving individual survey links. A total of 14 student trainee responses and 15 mentor responses was received, yielding an overall response rate of 90.6 percent. There were 13 paired sets of mentor-mentee responses (81.3 percent of the potential paired-response sets).

The second administration of the CREDIT URE assessment took place in spring 2017, after the first cohort of BUILD EXITO RLC students had been in their placements for approximately 1.5 years and the second cohort of RLC students had been in placements for approximately six months. For this second administration, the survey was emailed to 65 student-mentor pairs. Eleven mentors had more than one student, resulting in 65 students and 51 faculty receiving individual survey links. A total of 47 student trainee responses and 51 mentor responses was received (although not all mentors responded, 12 mentors responded for multiple mentees), resulting in an overall response rate of 75.4 percent, with 38 mentor-mentee completed paired responses (58.5 percent of total potential paired sets).

The third administration of the CREDIT URE assessment took place in spring 2018, after the second cohort of BUILD EXITO RLC student trainees had been in their placements for approximately 1.5 years and the third cohort had been in their placements for approximately six months (the first cohort graduated in spring 2017). The survey was emailed to 132 student-mentor pairs. Thirty mentors had more than one student, resulting in 132 students and 83 faculty receiving individual survey links. A total of 95 student trainee responses and 87 faculty responses was received, resulting in an overall response rate of 71.6 percent, with 67 mentor-mentee completed paired responses (50.8 percent of potential paired sets).

Response Rates

Response rates were maximized through group and individual email reminders as well as in-person reminders at required BUILD EXITO student trainee workshops. At two sites, paper surveys were administered to supplement the online links. Data were collected, entered, and stored in compliance with the BUILD EXITO IRB-approved protocol. No incentives were offered for survey completion; however, students were encouraged to complete all instruments as part of their participation in the program. The demographics of respondents were not significantly different from the BUILD EXITO cohorts as a whole.

Analysis

The analyses were designed to answer the four research questions presented above. First, for each given role, the percentage of student trainees with any responsibility for that role was determined, with corresponding percentages based on student self-report and mentor report.

Second, for respondents indicating any participation in a particular role (score higher than 0), the mean level of responsibility in that role was computed, again reflecting both student self-report and mentor report. The preceding analyses aggregated all responses higher than 0 across the three administrations. Third, using only the matching student-mentor response pairs from the third administration, the corresponding means for level of responsibility in each role as reported by students and by mentors were computed, as were percent agreement and kappa scores to assess degree of concordance. Finally, for the two cohorts of students assessed over time, the initial and final means for each role were compared. Means testing was done to assess the significance at $p < 0.05$ for any differences between student-mentor pairs and between initial and final administrations. All analyses were conducted using R statistical software, version 3.5, 2019.

Results

In which areas and to what extent do students make contributions to their RLCs?

To address research questions 1 and 2, contribution percentages for the three administrations were calculated (see Table 3). These results included all responses. Students and mentors most frequently reported students contributing to data curation, with close to 90 percent of responses indicating at least some involvement in that activity (see Table 3). Other roles endorsed at high rates (over 80 percent) by both students and mentors included investigation, formal analysis, visualization, and conceptualization. Other commonly reported roles, ranging between 60 and 75 percent of student and mentor respondents, included validation, project administration, methodology, and writing and reviewing/editing. The least common roles involved resources, software, supervision, and funding acquisition, although sizable percentages of students (approximately 25 to 50 percent) did engage in these activities.

Mean scores across all three administrations for student trainees and their mentors reporting at least some level of involvement in a role (students assigned a score greater than 0) also were computed. For example, for the 87.1 percent of students reporting involvement in investigation, the mean score was 2.3. On this scale of 1 (little responsibility) to 3 (primary responsibility), the roles receiving the highest average ratings from both students and mentors were investigation and data curation (means: 2.1–2.3). These higher ratings indicated not only that these were common roles but also roles for which student trainees had substantial responsibility (albeit "moderate" rather than "primary" responsibility). Other roles for which student trainees had considerable responsibility, with both student and mentor mean ratings at 1.9, included formal analysis and visualization. For roles with lower percentages of participating students, such as software, supervision, and funding acquisition, mentees and their mentors still reported notable responsibility, as indicated by student trainee and mentor means above 1.5.

How do contributions reported by students and by their mentors compare?

As shown in Table 3, the mean levels of engagement in the various research roles reported by student trainees and their mentors were generally aligned. Because the values reported in Table 3 were aggregated over all respondents, they did not provide a strong indicator of the agreement between mentees and mentors within mentoring pairs. Figure 1 compares

the ratings for the 67 matched student-mentor pairs completing the third administration of the CREDIT URE. This analysis included students from cohorts 2 and 3, and means were calculated using the full scale of no responsibility (0) to primary responsibility (3). The mean differences between mentees and their mentors illustrated where those groups had different impressions of students' levels of responsibility for certain tasks. Using the Wilcoxon rank-sum test, there were no roles in which differences between mentor and mentee responses were statistically significant ($p < 0.05$). Only two roles, validation and writing–original draft, had mean differences greater than 0.1. In both cases, students rated themselves slightly lower than the ratings of their mentors. Inter-rater reliability was an additional indicator of agreement. Percent agreements were calculated for each role, selecting this measure for ease of interpretability and appropriateness for the sample (well trained on the instrument; McHugh 2012). Weighted percent agreements ranged from 77.4 percent (funding acquisition) to 65.1 percent (validation).

Do student contributions change in type or degree over time?

Table 4 shows the initial and final means and standard deviations by role for the 33 student trainees who completed both administrations. Due to the small sample of mentors and the number of student RLC transfers in cohort 1, analyses for mentor responses in matched pairs for this cohort were not performed. Means presented in Table 4 use the full scale, with response options ranging from 0 to 3. As shown in Table 4, the overall mean across all roles increased significantly from 1.2 early in the placement to 1.5 later in the placement. Some level of increase in responsibility was observed for 12 of the 14 roles. Paired-sample t-tests were calculated for each role. The difference over time was statistically significant at $p < 0.05$ for resources (0.7 increase), formal analysis (0.6 increase), software (0.4 increase), and supervision (0.4 increase). The two roles that showed slight declines were data curation and validation, but these changes were not statistically significant.

Discussion

The initial use of the CREDIT URE provides preliminary evidence regarding the utility of this measure for assessing contributions to research made by undergraduate trainees in long-term research placements. First, the study demonstrated the feasibility of administering the CREDIT URE via online survey, with overall response rates averaging 73 percent for both student trainees and their research mentors, despite some variability across time points. Second, the instrument yielded meaningful distinctions between the different research roles as related to student participation and extent of student responsibility, as discussed further below. Third, the level of agreement between students and mentors on role engagement was good in the aggregate and very high in direct comparisons of pairs responding to the same case. Given the degree of corroboration provided by more experienced and objective research mentors, the credibility of student-reported data appears strong. Fourth, the CREDIT URE was sensitive enough to detect increases in student responsibility over time. Finally, analysis of data generated with the CREDIT URE suggests that the instrument can be employed for research and evaluation purposes, with implications for developing and improving research training programs.

The current study provides valuable insights regarding the nature and extent of undergraduate student research experiences in RLC placements. All roles represented on the CREDIT URE are considered important to the overall process of conducting and reporting research. Findings indicate that students in long-term RLC placements are gaining broad exposure to a variety of research activities. However, findings also suggest variability in the roles of students corresponding to their general levels of training and experience. For example, both student trainees and their mentors rated data curation and investigation as the most common roles and those with the highest level of student responsibility. Data curation includes data cleaning, maintenance, and management. Investigation includes performing experiments, including data collection. These particular roles may involve repetitive work guided by structured procedures and protocols. Thus, mentors can provide clear instruction and training, and the students can develop the relevant skills through practice, after which they can assume relatively high levels of responsibility.

A large number of roles was rated highly on participation but in the midrange regarding level of responsibility. This pattern may reflect activities in which students are receiving training with more active engagement and oversight by the research mentor. In other words, the student may have a part in the process, but the mentor still takes the lead and manages the process. These roles may represent the most common and productive domains for mentoring and learning. The nature of these roles suggests the possibility of this transfer of knowledge and experience through joint activity: conceptualization, formal analysis, visualization, validation, methodology, project administration, writing, etc. Of particular relevance and focus for the BUILD EXITO program is the level of student trainee engagement reported for writing, reviewing, and editing manuscripts. The program attempts to address the relatively low level of undergraduate and graduate student engagement in research dissemination through academic manuscripts (Garbati and Samuels 2013). Publications are a key research productivity metric across all phases of a research career. As a criterion for admission to graduate programs, publications may be of particular importance for underrepresented minorities (Keith-Spiegel, Tabachnick, and Spiegel 1994). A possible barrier to engaging students in writing manuscripts is that it is time intensive for mentors. Furthermore, the timing of a student's placement may not align with the time frame of a publication.

The research roles with the consistently lowest responsibility ratings include supervision, funding acquisition, and software (i.e., programming and software development). Undergraduate students are usually the least experienced and lowest ranking members of a research lab, so it is not surprising that they are not engaging in supervision of other lab members. As students gain more experience, they will be able to take on more responsibility for this role, perhaps becoming peer mentors or engaging with new students entering the same placement. Indeed, student trainees indicated increased responsibilities for this role over time. Some labs may have summer "shadow" opportunities for high school students, and undergraduates could be given responsibility for supervision of simple tasks with those students. It also is not surprising that tasks involving funding acquisition received low ratings. Funding acquisition is often the responsibility of the lab's principal investigator and involves a high level of expertise. In about one quarter of EXITO RLCs, the student's mentor is not the principal investigator but is a postdoctoral student or another team member who does not have direct responsibility for funding acquisition. Thus, students may have

limited exposure to the process of identifying funding sources and writing grants. Even if students are engaged in such an effort by virtue of conducting a literature review, generating pilot data, or compiling appendices, they may not have a high level of responsibility for the ultimate grant submission.

One of the most important findings of the study concerns the anticipated increases in student trainee responsibilities during the one year between the two survey administrations. Consistent with other research showing that students develop a greater understanding of the research process with longer placement times (Thiry et al. 2012), results indicate that students increased their responsibilities as they gained exposure and experience. A statistically significant gain in responsibility was observed across the two timepoints when all the roles were averaged. Students reported gains in responsibility, from early in their placements to later in their placements, for 12 of the 14 roles. Of note, several of the largest and statistically significant increases were observed in roles that initially had lower levels of engagement, such as resources, software, and supervision. Another statistically significant increase was in the role of formal analysis, and there were notable gains in some other areas that may encourage more active mentoring (e.g., visualization, writing, project administration). Not surprisingly, the most highly rated roles, investigation and data curation, had minimal change over time. This pattern of results may reflect the natural progression of activities in the research placement, as suggested above. For example, students may start with the routine roles of investigation and data curation, improve in areas that involve more coaching and mentoring such as formal analysis, and ultimately assume greater responsibility for supervising others and managing lab resources.

Another factor that may explain the changes in ratings over time is that students may develop a better sense of the respective roles. For example, although mentor and mentee ratings were remarkably congruent, the few observed differences were consistent with other work comparing student and mentor ratings within the URE setting (Cox and Andriot 2009). When asked informally, BUILD EXITO RLC mentors theorized that score differences between mentors and students could be partially attributed to the fact that mentors are likely to have a higher level of familiarity with the roles. To address this, the CREDIT URE could be introduced early in the placement and mentoring relationship so students have time to think more about their roles prior to CREDIT URE administration.

Limitations

Limitations of this work included the sample size, student lab transitions between data collection points, and student level of understanding of lab role definitions. The study had a relatively small sample size, particularly for the pilot administration sample. As the BUILD EXITO program continues, larger sample sizes will be obtained, providing greater confidence in results. In future work administering the CREDIT URE, other incentives to increase response rates will be considered, although the study's response rates were moderate to good. Preliminary analyses suggested that there were minimal demographic differences between responders and nonresponders. Second, approximately half of all cohort 1 students switched lab placements between the two pretest/posttest time points. This reflected the program implementation period, and the pattern was not observed in latter

cohorts. This limited the conclusions that could be drawn from cohort 1 analyses, as students might have engaged in different sets of activities across different labs. Finally, increases in mean scores across years could be due to an increased understanding of role definitions by students, and future work should consider examining the possible role of this increased understanding in mean score changes.

Conclusion

Analyses indicate that this novel survey tool, the CREDIT URE, is appropriate for both formative and summative measurement. Unlike measures that rely on self-report of research skills, the CREDIT URE measures the actual degree of student participation in a continuum of tasks considered essential to the research process. Same-construct ratings from students' mentors provide a check to self-report biases, and administration at multiple time points allows students, mentors, and program staff to see student growth. Introduction to the CREDIT taxonomy provides an opportunity for students to learn about the breadth of research roles and provides an opportunity for mentors to assess the extent of tasks assigned to their students and realign task assignment if necessary. In addition, the CREDIT URE has potential for use with students at multiple levels of career development, including undergraduate, graduate, and postdoctorate.

Next steps for this work include administrations of the CREDIT URE to additional cohorts of BUILD EXITO students. Larger samples will allow for subgroup analyses and comparisons, such as comparing students who transfer between RLCs versus those who stay in the same RLC for the full 18 months. Future work also will take into account demographic variables such as gender. Other variables such as prior research experience will be assessed and integrated into analyses, building on previous work that examines outcomes for novice and experienced undergraduate researchers (Thiry et al. 2012). An attempt will be made to determine the predictive validity of the CREDIT URE by examining BUILD EXITO student outcomes, including application and matriculation to graduate school. Evaluation of BUILD EXITO RLCs will be expanded to include qualitative evaluation, using the CREDIT URE as part of a mixed-methods approach, with student trainees and mentors providing examples of tasks for each level of the taxonomy.

Acknowledgments

Biography

Matt Honoré is research associate for BUILD EXITO, an NIH-funded undergraduate research training program. Prior to EXITO, Honoré supported the administration of several federal AmeriCorps programs in Oregon. He received a master's degree in public administration with a specialization in nonprofit management from Portland State University (PSU).

Thomas E. Keller is the Duncan and Cindy Campbell Professor for Children, Youth, and Families at the PSU School of Social Work and co-investigator of the BUILD EXITO project. His research interests include the development and influence of mentoring relationships, the evaluation of mentoring program innovations and enhancements, and the mentoring of undergraduates historically underrepresented in biomedical and health sciences.

Jen Lindwall is a doctoral student in applied developmental psychology at PSU, where she examines factors that support minority students in STEM disciplines. After spending a decade working in the mentoring field, she transitioned to higher education and currently works as a research assistant for BUILD EXITO.

Rachel Crist works as a researcher at Oregon Health and Science University (OHSU) and holds an MS in communication studies from PSU. Her work focuses on engaging participants and improving participant experience, equity, and access to health care as well as higher education.

Leslie Bienen is a veterinarian and professor at the OHSU–PSU School of Public Health, as well as the pilot project coordinator for BUILD EXITO.

Adrienne Zell is the director of the OSHU Evaluation Core, an assistant professor at the OHSU–PSU School of Public Health, and assistant director at the Oregon Clinical and Translational Research Institute. Zell and her team lead evaluation for BUILD EXITO. She holds a PhD from the University of Chicago.

References

Brand Amy, Allen Liz, Altman Micah, Hlava Marjorie, and Scott Jo. 2015. "Beyond Authorship: Attribution, Contribution, Collaboration, and Credit." Learned Publishing 28: 151–155. doi: 10.1087/20150211

Chang Mitchell J., Sharkness Jessica, Hurtado Sylvia, and Newman Christopher B.. 2014. "What Matters in College for Retaining Aspiring Scientists and Engineers from Underrepresented Racial Groups." Journal of Research in Science Teaching 51: 555–580. doi: 10.1002/tea.21146

Cox Monica F., and Andriot Angie. 2009. "Mentor and Undergraduate Student Comparisons of Students' Research Skills." Journal of STEM Education: Innovations and Research 10(1/2): 31–39.

Dyer-Barr Raina. 2014. "Research to Practice: Identifying Best Practices for STEM Intervention Programs for URMs." Quality Approaches in Higher Education 5(1): 19–25.

Eagan M. Kevin Jr., Hurtado Sylvia, Chang Mitchell J., Garcia Gina A., Herrera Felisha A., and Garibay Juan 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 [PubMed: 25190821]

Feldman Allan, Divoll Kent A., and Rogan-Klyve Allyson. 2013. "Becoming Researchers: The Participation of Undergraduate and Graduate Students in Scientific Research Groups." Science Education 97: 218–243. doi: 10.1002/sce.21051

Fisher Robert J., and Katz James E.. 2000. "Social-Desirability Bias and the Validity of Self-Reported Values." Psychology & Marketing 17: 105–120. doi: 10.1002/(sici)1520-6793(200002)17:2<105::aid-mar3>3.0.co;2-9

Garbati Jordana, and Samuels Boba. 2013. "Publishing in Educational Research Journals: Are Graduate Students Participating?" Journal of Scholarly Publishing 44: 355–372. doi: 10.3138/jsp.44-4-004

Gregerman Sandra R., Lerner Jennifer S., William von Hippel John Jonides, and Nagda Biren A.. 1998. "Undergraduate Student-Faculty Research Partnerships Affect Student Retention." Review of Higher Education 22: 55–72. doi: 10.1353/rhe.1998.0016

Hurtado Sylvia, Cabrera Nolan L., Lin Monica H., Arellano Lucy, and Espinosa Lorelle L.. 2009. "Diversifying Science: Underrepresented Student Experiences in Structured Research Programs." Research in Higher Education 50: 189–214. doi: 10.1007/s11162-008-9114-7 [PubMed: 23503690]

Keith-Spiegel Patricia, Tabachnick Barbara G., and Spiegel Gary B.. 1994. "When Demand Exceeds Supply: Second-Order Criteria Used by Graduate School Selection Committees." Teaching of Psychology 21: 79–81. doi: 10.1207/s15328023top2102_3

Keller Thomas E., Logan Kay, Lindwall Jennifer, and Beals Caitlyn. 2017. "Peer Mentoring for Undergraduates in a Research-Focused Diversity Initiative." Metropolitan Universities 28(3): 50. doi: 10.18060/21542 [PubMed: 29398880]

Linn Marcia C., Palmer Erin, Baranger Anne, Gerard Elizabeth, and Stone Elisa. 2015. "Undergraduate Research Experiences: Impacts and Opportunities." Science 347: 1261757. doi: 10.1126/science.1261757 [PubMed: 25657254]

Lopatto David. 2007. "Undergraduate Research Experiences Support Science Career Decisions and Active Learning." CBE Life Sciences Education 6: 297–306. doi: 10.1187/cbe.07-06-0039 [PubMed: 18056301]

Mathews Andrew, and Bradle Brendan. 1983. "Mood and the Self-Reference Bias in Recall." Behaviour Research and Therapy 21: 233–239. doi: 10.1016/0005-7967(83)90204-8 [PubMed: 6615388]

McHugh Mary L. 2012. "Interrater Reliability: The Kappa Statistic." Biochemia Medica 22: 276–282. doi: 10.11613/bm.2012.031 [PubMed: 23092060]

National Center for Education Statistics (NCES). 2015. "Post-secondary Attainment: Differences by Socioeconomic Status," in The Condition of Education 2015 (NCES 2015–144). Accessed September 15, 2020. https://nces.ed.gov/programs/coe/indicator_tva.asp

National Institutes of Health (NIH). 2019. "Notice of NIH's Interest in Diversity" (NOT-OD-20–03). Accessed September 15, 2020. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-20-031.html

Richardson Dawn M., Keller Thomas E., Wolf De'Sha S., Zell Adrienne, Morris Cynthia, and Crespo Carlos J.. 2017. "BUILD EXITO: A Multi-Level Intervention to Support Diversity in Health-Focused Research." BMC Proceedings 11: 19. doi: 10.1186/s12919-017-0080-y [PubMed: 29375661]

Russell Susan H., Hancock Mary P., and McCullough James. 2007. "Benefits of Undergraduate Research Experiences." Science 316: 548–549. doi: 10.1126/science.1140384 [PubMed: 17463273]

Seymour Elaine, Hunter Anne-Barrie, Laursen Sandra L., and Deantoni Tracee. 2004. "Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study." Science Education 88: 493–534. doi: 10.1002/sce.10131

Shortlidge Erin E., Bangera Gita, and Brownell Sara E.. 2015. "Faculty Perspectives on Developing and Teaching Course-Based Undergraduate Research Experiences." BioScience 66: 54–62. doi: 10.1093/biosci/biv167

Smith Elise, and Williams-Jones Bryn. 2012. "Authorship and Responsibility in Health Sciences Research: A Review of Procedures for Fairly Allocating Authorship in Multi-Author Studies." Science and Engineering Ethics 18: 199–212. doi: 10.1007/s11948-011-9263-5 [PubMed: 21312000]

Thiry Heather, Laursen Sandra L., and Hunter Anne-Barrie. 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: 357–388. doi: 10.1080/00221546.2011.11777209

Thiry Heather, Weston Timothy J., Laursen Sandra L., and Hunter Anne-Barrie. 2012. "The Benefits of Multi-Year Research Experiences: Differences in Novice and Experienced Students' Reported

Gains from Undergraduate Research." CBE–Life Sciences Education 11: 260–272. doi: 10.1187/cbe.11-11-0098 [PubMed: 22949423]

Tsui Lisa. 2007. "Effective Strategies to Increase Diversity in STEM Fields: A Review of the Research Literature." Journal of Negro Education 76: 555–581.

Valantine Hannah A., and Collins Francis S.. 2015. "National Institutes of Health Addresses the Science of Diversity." Proceedings of the National Academy of Sciences 112: 12240–12242. doi: 10.1073/pnas.1515612112

Weston Timothy J., and Laursen Sandra L.. 2015. "The Undergraduate Research Student Self-Assessment (URSSA): Validation for Use in Program Evaluation." CBE–Life Sciences Education 14(3): ar33. doi: 10.1187/cbe.14-11-0206 [PubMed: 26250563]
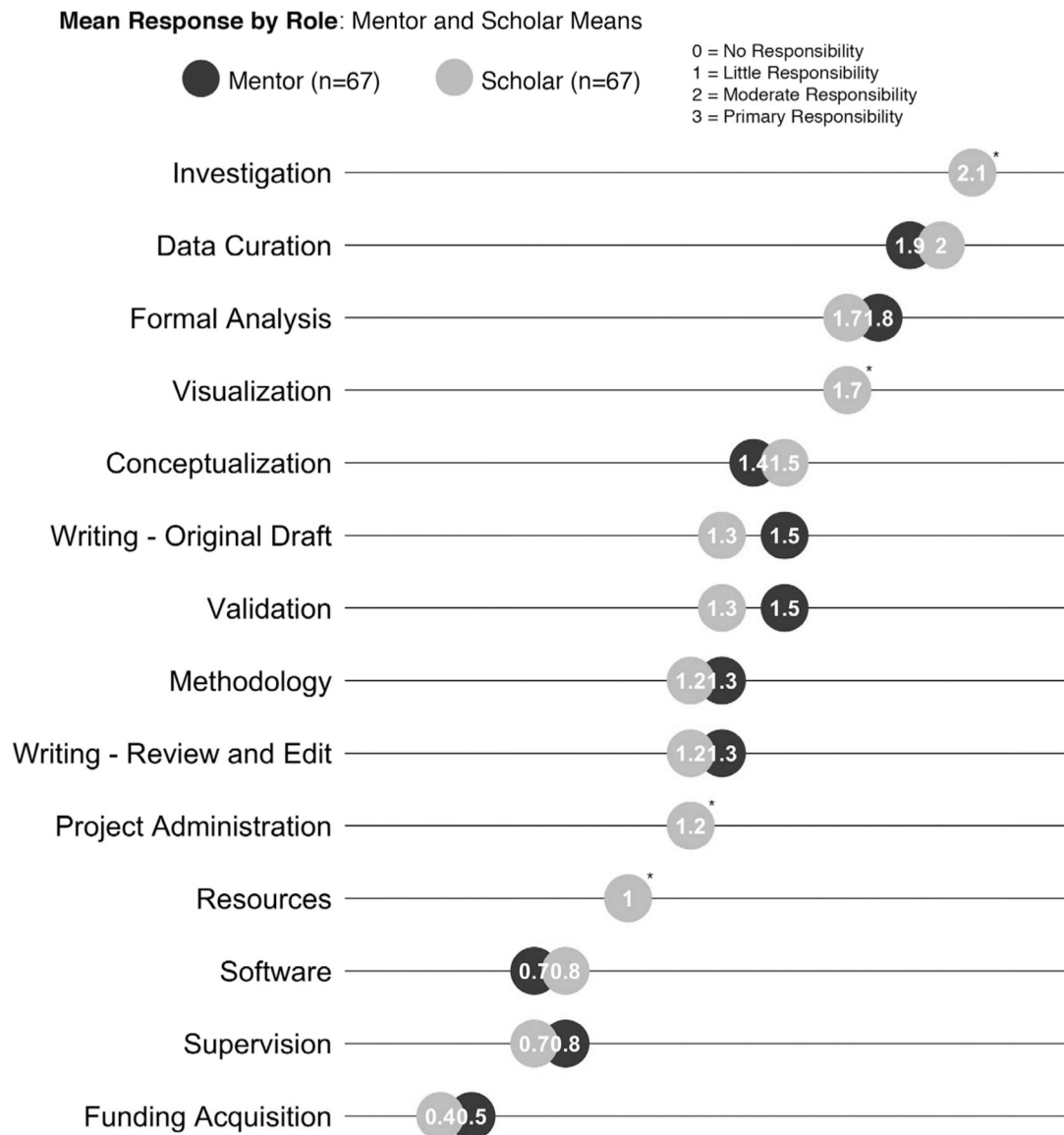
FIGURE 1. Mean Response by Role: Mentor and Scholar Means
*Denotes Mentor and Scholar means overlap when rounded to the 10th decimal point

TABLE 1.

The Open-Source CRediT Taxonomy

Role	Definition
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims
Data curation	Management activities to annotate (produce metadata), clean data, and maintain research data (including software code, where necessary for interpreting the data itself) for initial use and later reuse
Formal analysis	Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data
Funding acquisition	Acquisition of the financial support for the project leading to publication
Investigation	Conduct of a research and investigation process, specifically performing the experiments or data/evidence collection
Methodology	Development or design of methodology; creation of models
Project administration	Management and coordination responsibility for research activity planning and execution
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, and other analysis tools
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team
Validation	Verification, whether as a part of the activity or separately, of the overall replication/reproducibility of results/experiments and other research outputs
Visualization	Preparation, creation, and/or presentation of the published work, specifically visualization/data presentation
Writing—original draft	Preparation, creation, and/or presentation of the published work, specifically writing the initial draft (including substantive translation)
Writing— reviewing and editing	Preparation, creation, and/or presentation of the published work by members of original research group, specifically critical review, commentary, or revision, including pre- and post-publication stages

Note: CRediT = Contributor Roles Taxonomy

TABLE 2.

CREDIT URE Response Rate: Pilot, Second, and Third Administrations

Respondents/sample	Pilot administration (cohort 1)	Second administration (cohorts 1 and 2)	Third administration (cohorts 2 and 3)	Pilot and second administration (cohort 1: within subjects, pretest/posttest)	Second and third administration (cohort 2: within subjects, pretest/posttest)
			completed survey/total *n* sent (response rate)		
Mentors[a]	15/16 (91.7%)	51/65 (78.5%)	87/132 (66.0%)	NA[b]	15/22 (68.2%)
Students	14/16 (87.5%)	47/65 (72.3%)	95/132 (72.0%)	9/14 (64.3%)	24/37 (64.9%)
Cohort 1	14/16 (87.5%)	10/14 (71.4%)	–	9/14 (64.3%)	–
Cohort 2	–	37/51 (72.5%)	32/43 (74.4%)	–	24/37 (64.9%)
Cohort 3	–	–	63/89 (70.8%)	–	–
Student pretest/posttest responses	–	–	24/37 (64.9%)	–	–
Student-mentor paired responses	13/16 (81.3%)	38/65 (58.5%)	67/132 (50.8%)	–	–

[a]Total number of mentor surveys distributed; some mentors had more than one student mentee.

[b]Due to a high degree of RLC turnover in cohort 1, mentor/mentee pretest/posttest response rates were not calculated.

TABLE 3.

Student and Mentor CREDIT URE Means for All Respondents Reporting Any Involvement

	Students				Mentors			
	% reporting any involvement	Mean level of responsibility	SD	N	% reporting any involvement	Mean level of responsibility	SD	N
Investigation	87.1	2.3	0.7	135	86.4	2.3	0.7	127
Data curation	89.4	2.3	0.8	135	87.7	2.1	0.7	128
Formal analysis	81.2	1.9	0.7	125	81.8	1.9	0.7	121
Visualization	80.8	1.9	0.8	122	84.9	1.9	0.7	124
Conceptualization	87.6	1.7	0.6	134	83.1	1.6	0.7	123
Writing—original draft	71.1	1.8	0.8	106	69.4	1.9	0.7	102
Validation	74.0	1.7	0.7	111	70.9	1.8	0.8	105
Methodology	72.4	1.7	0.7	110	71.9	1.7	0.8	105
Writing—reviewing and editing	69.5	1.8	0.7	105	66.7	1.8	0.7	98
Project administration	73.4	1.8	0.7	113	61.6	1.8	0.7	90
Resources	57.7	1.7	0.8	86	50.7	1.6	0.7	75
Software	44.5	1.7	0.7	69	32.9	1.7	0.7	48
Supervision	45.8	1.6	0.7	70	38.4	1.6	0.7	56
Funding acquisition	29.3	1.6	0.7	44	24.2	1.7	0.8	36

TABLE 4.

Pretest and Posttest Means, Student Trainees, Differences within Subjects

Role	Within subject means			N
	Time 1 Spring RLC placement Year 1	Time 2 Spring RLC placement Year 2	Difference of means	
	Mean (SD)	Mean (SD)		
Investigation	2.12 (0.99)	2.18 (0.81)	0.06	33
Data curation	2.03 (1.15)	1.97 (0.98)	−0.06	29
Formal analysis	1.36 (1.03)	1.91 (0.88)	0.55^a	33
Visualization	1.52 (1.12)	1.88 (0.78)	0.36	33
Conceptualization	1.52 (0.83)	1.64 (0.60)	0.12	33
Writing—original draft	1.25 (1.05)	1.66 (1.07)	0.41	32
Validation	1.34 (1.08)	1.31 (0.81)	−0.03	29
Methodology	1.24 (0.94)	1.39 (0.93)	0.15	33
Writing—review and edit	1.16 (0.88)	1.38 (1.10)	0.22	32
Project administration	1.16 (0.99)	1.59 (0.84)	0.43	32
Resources	0.65 (0.88)	1.35 (0.98)	0.70^a	31
Software	0.48 (0.48)	0.91 (1.04)	0.43^a	33
Supervision	0.52 (0.52)	0.90 (0.91)	0.38^a	31
Funding acquisition	0.39 (0.70)	0.58 (0.94)	0.19	33
All roles	1.19 (1.07)	1.47 (1.00)	0.28^a	

Note: N = 33

[a] Statistically significant at $p < 0.05$.