What We Know and What We Need to Know about Undergraduate Research

Heather Haeger, John E. Banks, California State University, Monterey Bay
Camille Smith, Monique Armstrong-Land, California State University, Monterey Bay

Abstract
To assess what questions are already answered and what still needs to be discovered about the high-impact practice of undergraduate research (UR), the authors conducted a mixed-methods study, including a systematic analysis of literature that assessed the impact of UR, and interviewed faculty and administrators actively engaged in UR. Findings demonstrated that most studies on UR have focused on STEM fields and student outcomes. Fewer studies have examined other disciplines or other outcomes such as the impact of UR on faculty or institution. Despite ample research that demonstrates outcomes associated with UR, more work is needed to establish a causal relationship between UR and these outcomes, to diversify the topics and scope of scholarship on UR, and to demonstrate the far-ranging impacts of UR.

Keywords: assessment, impacts, literature review, undergraduate research outcomes

doi: 10.18833/spur/3/4/4

Methods
To review what research had already established about UR and to identify gaps in the current body of knowledge, a mixed-methods approach was taken that included interviews, focus groups, and a systematic review of the literature (Grant and Booth 2009). Interviews and focus groups were conducted (N = 69) in 2016–2017 with leaders in UR (see Figure 1). Participants were recruited from leaders of
After the interviews and focus groups were completed, the researchers collected the sorted sticky notes and the notes taken by the research team during the discussion. These were compiled into a list of all of the questions that were unanswered or outcomes that had not been well established through rigorous research. Inductive coding was used to synthesize the questions into four main themes: (a) fundamental questions that need to be addressed, refined, or replicated; (b) questions about student experiences and outcomes; (c) questions about faculty and mentor experiences and outcomes; and (d) questions about the broader impacts of UR.

Questions were coded as fundamental if they were brought up in the focus groups but also were addressed in prior studies in the literature review. These were questions or topics that needed more rigorous research or more standardized measurements despite their place in previous studies. The three other codes were based on the focus of the outcomes: student outcomes, faculty outcomes, and broader institutional and societal impacts. The structure of this analysis also was vetted by CUR leadership, presented to the CUR Board of Directors for its review, and refined based on participant confirmation and feedback (Jones, Torres, and Arminio 2013).

In addition to interviews, the authors conducted a systematic review of journal articles about undergraduate research (Grant and Booth 2009). This review focused on the assessment of undergraduate research, starting with 263 articles selected from the CUR annotated bibliography on assessing undergraduate research experiences or programs (CUR n.d.). An additional 102 articles were found by searching the Education Resources Information Center (ERIC) for articles using the key term undergraduate research; these articles directly assessed the impact of undergraduate research from 2000 to present. Articles also were limited to higher education, the English language, and academic journals. Duplicate articles from the database were eliminated, and only peer-reviewed articles and studies in which the assessment of undergraduate research was the primary focus were selected. A specific search was performed for articles that qualified for the WWC, but the search returned no results. From the 365 articles produced by the search, 286 articles were selected for further review and analysis based on the strengths of their research design and their relationship to the key questions identified by the interviews and focus groups.

All articles were reviewed by both an undergraduate research assistant and the lead author. Coding was consistent for 94 percent of the articles, and any inconsistencies were discussed and resolved. Articles were coded for focus; type of UR (mentored or course-based undergraduate research experience, or CURE); and outcome and...
What We Know and What We Need to Know

methodology, that is, indirect measures or direct measures of student learning or development, literature reviews, and potential qualification for the WWC (see Figure 2). Indirect measures included surveys, focus groups, interviews, and student reflections. Direct measures included testing specific content knowledge, GPA, retention or persistence in field and at institution, and graduation rates. Scoring of student products was designated a direct measure if the article clearly outlined methods for standardization (e.g., using blind review of students’ research posters and scoring them using an established and tested rubric).

Limitations
The systematic review of previous literature included a large sample of UR-related research but did not include every article on UR. Counts and percentages represented trends in the literature but should not be taken as a definitive count of all research published on UR (Grant and Booth 2009). Some articles that met the selection criteria ($n = 24$) were not readily accessible and did not provide enough detail in their abstracts to be fully categorized. This was particularly a problem for determining the methodology of the research. Those articles were excluded from the reports on methodology. Additionally, WWC standards were very specific and difficult to meet for research on educational practices like UR. WWC categorization was based on the standards outlined in 2017. Only articles that provided enough detail about their methods could be considered for this category. Trying to determine whether or not they met the standards also was a cumbersome process, and only articles with abstracts that suggested that the research used an experimental design or conducted a post-hoc analysis to create a quasi-experimental design were reviewed for WWC categorization.

Results
A substantial body of literature has explored the role of UR and its impact. To survey the composition of this body of literature, articles were categorized according to their focus, type of UR, and outcome (see Table 1). Articles related to student outcomes were further analyzed based on methodology (see Table 2).

A large number of articles assessing student outcomes focused on out-of-class, mentored UR (43 percent), although a significant number also assessed course-based UR (27 percent). The remaining articles were not about a specific type of undergraduate research (i.e., they focused on the topic or assessment of UR generally, were literature reviews, or did not provide enough detail about the type

Note: These categories were not mutually exclusive; therefore, percentages in this article’s tables do not add up to 100 percent. For example, an article about an engineering UR experience would be coded as both discipline specific and STEM. CURE = course-based undergraduate research experience. WWC = What Works Clearinghouse.
A majority of articles focused on STEM students or discipline-specific programs. Of the discipline-specific articles, 66 percent \((n = 88)\) also covered STEM discipline, and 34 percent \((n = 46)\) were about non-STEM disciplines. Very few articles \((n = 11)\) were focused explicitly on interdisciplinary programs.

Unsurprisingly, the majority (75 percent) of the research on the impact of UR was focused on outcomes for students, with only 13 percent considering faculty- or mentor-related outcomes (see Table 1). Of the articles that addressed faculty or mentor outcomes, the majority also considered student outcomes \((27 \text{ of } 36)\). A number of articles were descriptions of programs, descriptions of assessment methods, or literature reviews and were not coded as either outcome.

Although other articles referenced issues of diversity briefly, only 12 percent \((n = 35)\) of the articles explicitly focused on issues of diversity, inclusion, or representation of traditionally underrepresented populations.

The majority (64 percent) of articles assessing the impact of UR on student outcomes utilized indirect measures of student learning and development, including surveys, interviews, focus groups, and written reflections (see Table 2). A smaller proportion (30 percent) used more direct measures of learning, including GPA, retention at institution or in discipline, and tests of specific content knowledge, or they used a mix of direct and indirect measures. The remaining 5 percent of the articles were coded as literature reviews, because they did not directly test an outcome but rather reviewed previous literature on student outcomes.

The research on UR is subject to the limitations of educational and social science research in that it is frequently descriptive or relational, and few studies are designed that allow establishment of causality. Only 4 percent of the studies met the criteria, with reservations, for research on best practices laid out by the WWC; none of these articles met WWC criteria without reservation. A larger proportion...
of the articles on CUREs used more concrete or direct measures with an experimental or quasi-experimental design, a methodology approved by the WWC (four out of five WWC-eligible articles were about CUREs, one was about mentored UR, and two were about how to use methodology approved by the WWC).

The Big Questions

In addition to reviewing published literature, the authors consulted with leaders in UR about the most pressing issues to be addressed in UR research. The interviews and focus groups identified four main areas of UR that needed further research.

1. Fundamentals: the need for systematic and consistent standards to define, track, and measure undergraduate research participation.

2. Student outcomes: rigorous research focusing on the impact of research participation on students, with particular attention to the type, quality, and frequency of the UR experience.

3. Faculty outcomes: research regarding the impact of mentoring students in research on faculty, postdoctoral researchers, and graduate students.

4. Broader impacts: research on the larger benefits of undergraduate research.

The first three categories were brought up in each of the interviews and focus groups. Only the focus groups at the CUR Executive Board meeting and AAC&U discussed the need for study of the broader impacts of research, beyond student and faculty outcomes. Despite the designation of all these areas of research as those needing further work, there had been prior scholarship for each category; this should be kept under consideration when examining directions for future research. This section will present the needs for research discussed by interview and focus group participants and how these needs relate to any prior research to illustrate the limits of what is known and describe what remains to be explored.

Fundamentals

The first area of need is related to fundamental questions to be addressed, refined, or replicated. It includes the use of consistent definitions of UR and effective tracking of participation (Childress 2015; Crowe and Brakke 2019). The creation, testing, and use of validated instruments to measure quality and intensity of participation as well as student learning and development also are needed. National calls have stressed the need for research that utilizes experimental or quasi-experimental design and samples from multiple institutions, which would qualify for WWC. Past research that has utilized experimental or quasi-experimental methods (Corwin, Graham, and Dolan 2015; Nagda et al. 1998; Pender et al. 2010; Russell and Weaver 2011) should be updated and replicated.

Student Outcomes

Despite the demonstrated benefit of high-impact practices like UR to first-generation and underrepresented minority students, traditional patterns of unequal participation persist in UR (Haeger, BrckaLorenz, and Webber 2015; Kinzie et al. 2008). As indicated in the NASEM (2017) report, a better understanding is required regarding issues of equity and social justice related to research participation by traditionally underrepresented students.

To add to the established understanding of the impact of UR participation on academic performance and achievement, further work should evaluate the more complex and intangible benefits of UR, such as intellectual engagement, lifelong learning, and refinement of career aspirations (Hathaway, Nagda, and Gregerman 2002; Hunter, Laursen, and Seymour 2007; Kuh 2008; Kuh et al. 2010; Russell, Hancock, and McCullough 2007). Furthermore, research should be performed regarding how the details of the UR experience affect student outcomes. In particular, there is a need for more scholarship that explores the effects of teaching practices on the quality of course-based UR, as well as how student outcomes are affected by the intensity and duration of UR experiences or participation in multiple high-impact practices (Haeger and Fresquez 2016; Hanauer and Dolan 2014). Finally, there is a call for more research on the long-term impact of UR participation on students (Bauer and Bennett 2003; Johnson Schmitz and Havholm 2015; Zydney et al. 2002).

Faculty Outcomes

Questions about faculty and mentor experiences and outcomes also are crucial to understanding the impact of UR. First, more explicit cost-benefit analyses relating UR to faculty productivity and progress to promotion and tenure can bolster understanding of how institutions could better support faculty. As more colleges and universities focus on scaling up UR opportunities and incorporating UR into the classroom, research should include cost-benefit analyses of course-based UR (Auchincloss et al. 2014; Corwin et al. 2015; Laursen, Seymour, and Hunter 2012).

Furthermore, additional research is needed to better understand patterns of faculty participation in UR. Questions such as which faculty engage in UR and why (Eagan et al. 2011), as well as cost-benefits analyses related to tenure and promotion, can elucidate ways to incentivize faculty participation in UR. Future research should extend cost-benefit studies of faculty participation to effects on graduate and postdoctoral mentors as well (Dolan and Johnson 2009).

Broader Impacts

Finally, participants in focus groups and interviews frequently discussed the need for research that addresses the impact of UR beyond student learning and explores
its broader impacts. In particular, beyond the benefits for individual students or faculty, there is a need to better understand the impacts of UR on campuses: Do campuses with higher levels of UR participation receive more extramural funding, or contribute more to knowledge creation? Furthermore, how are the local and regional job markets affected (e.g., does UR participation create a more informed and productive workforce or employees that are better at problem solving)? Although these are ambitious and large-scale endeavors, more resources should be devoted to exploring how various levels of investment and participation in UR on different types of campuses influence campus and community outcomes to paint a more holistic portrait of the impacts of UR.

Conclusion

The substantial body of literature on UR developed over the past two decades has demonstrated numerous benefits to students (Bauer and Bennett 2003; Johnson Schmitz and Havholm 2015; Zydny et al. 2002) and to faculty (Dolan and Johnson 2009; Eagan et al. 2011). Despite this, more work is needed to establish a causal relationship between UR and these outcomes, to diversify the topics and scope of research on UR, and to demonstrate the far-ranging impacts of UR.

A common theme emerging from the analyses of the literature, focus groups, and interview results was the need for more rigorous research on the impacts of UR. Anecdotally, it is clear to those who have mentored students in UR and to undergraduates who have participated themselves that UR can have tremendous benefits for students and mentors. Quantifying these effects is another matter; to date, the majority of research measuring the impact of UR relies on indirect measures or correlations between outcomes and participation. Both NASEM and WWC have highlighted the need for more rigorous research on UR and for replication of foundational studies. Additionally, most scholarship on UR to date has been based on a single campus or single program, and further research is needed to assess UR across campuses and systems. Conducting multi-campus, experimental, or quasi-experimental research on UR is challenging because of ethical and practical barriers to randomly assigning students to participate in UR experiences. Some research studies have addressed this by creating a lottery for participation in UR (Nagda et al., 1998) or conducting post-hoc analyses such as propensity score matching to create control groups (Haeger and Fresquez 2016). Research on course-based undergraduate research also has incorporated multiple sections of the same class to approximate random assignment to treatments and create a control group (Brownell et al. 2013; Corwin et al. 2015). Future research should continue to explore innovative ways to rigorously assess the impacts of UR across multiple platforms.

Despite the importance of UR as a high-impact practice with disproportionately high gains for first-generation undergraduate students and students of color, underserved populations and low-income and transfer students are still less likely to participate in UR than other groups (Haeger et al. 2015). NASEM (2017) has highlighted the social justice issue of unequal representation in UR as a key area of focus for further study. The problem of unequal participation is compounded by the need for more focus on diversity and inclusion in published research. Although some of the surveyed articles referenced issues of diversity briefly, only 12 percent (n = 35) of the articles explicitly focused on issues of diversity, inclusion, or representation of traditionally underrepresented populations.

In addition to addressing issues of diversity in student participants, further research is needed that addresses the diversity of fields encompassed by UR. The fact that most of the articles surveyed focused on STEM in general or a STEM discipline reflects the historical predominance of UR in STEM fields. Scholarship exploring UR beyond STEM, including interdisciplinary fields, is desperately needed.

Apart from specific calls to explore the increasingly nuanced ways in which UR may impact students, faculty, and institutions, the focus groups and interviews revealed a need to find better ways of measuring the less tangible impacts of UR. As a transformative experience for students, UR has the potential to change how students think, how the careers they see as possible for themselves can multiply, and how they interact with and question the world around them. Similarly, faculty engaged in UR may experience marked changes in how they interact with student researchers, with attendant consequences for their own personal and professional development, career trajectories, and research accomplishments. Finally, it is important to consider how UR impacts educational institutions, communities, and ultimately society as a whole. The challenge to future researchers is to find reliable ways to define, measure, and test the full breadth of the impact of undergraduate research in all of its forms and complexity.

References


Heather Haeger is the educational research associate at the Undergraduate Research Opportunity Center (UROC) at California State University, Monterey Bay (CSUMB). She conducts research on educational equity in high-impact practices and the transition to graduate education. Her research is used to inform programmatic interventions aimed at engaging traditionally underserved populations. Haeger also serves as the CUR assessment and research coordinator. She surveys current literature on undergraduate research, develops tools for use by CUR members to conduct their own research and assessment, and studies the role of faculty mentoring of undergraduate research in tenure and promotion. Haeger received her PhD in education policy, studies, and practices from the University of Arizona. From 2011–2014, she was an assistant research faculty in the Center for Postsecondary Research at Indiana University, where she designed, tested, and administered surveys on student engagement, including the National Survey of Student Engagement (NSSE) and the Law School Survey of Student Engagement (LSSSE).

John E. Banks is director of UROC at CSUMB. He earned an MS in applied mathematics at the University of Southern California and a PhD in zoology at the University of Washington, Seattle. He spent 16 years as a faculty member and administrator at the Tacoma campus of the University of Washington, where he was professor of environmental science and served as director of international programs and director of undergraduate education. Banks trained as a quantitative ecologist and has published more than 75 scientific articles, book chapters, and technical reports—several with undergraduate coauthors—exploring biological control and predator/prey population dynamics in agroecosystems, ecotoxicology, and the conservation of biological diversity in natural and managed systems in both temperate and tropical environments.

Camille Smith is a collaborative health and human services major at CSUMB. She is currently a UROC researcher under the mentorship of Heather Haeger, assessing the impact of various interventions on the success of students engaged in undergraduate research and identifying perceived barriers to student engagement in undergraduate research.

Monique Armstrong-Land is a student in humanities and communication at CSUMB. She is currently a UROC ambassador and under the mentorship of Haeger has conducted research to identify barriers to the engagement of traditionally underrepresented students in undergraduate research.