Methods Matter: Contrasting Undergraduate Research Experience Outcomes based on Surveys and Interview Methods

Angela R. Bielefeldt, Lupita Montoya
Dept. Civil, Environmental, & Architectural Engineering
University of Colorado Boulder
Boulder, CO, USA
Angela.Bielefeldt@colorado.edu

Greg Rulifson
Division of Engineering, Design, and Society
Colorado School of Mines
Golden, CO, USA

***Abstract*— Undergraduate students who engage in research can realize a variety of personal and professional benefits. Studies of undergraduate research experiences in engineering have largely relied on survey methods, which can be subject to positive response bias due to acquiescence and/or social desirability. This research compared two convenience sources of information to determine the impacts of undergraduate research experiences for engineering students. Survey findings from 8-years assessing a summer REU site focused on environmental engineering (n=85) tended to be overwhelmingly positive based on Likert-type responses to pre and post survey questions. Open-ended responses were enlightening in regards to the most valuable aspect of the summer REU experience for the students. The second data source were yearly student interviews not explicitly targeted to study research experiences; nine engineering students discussed their research experiences. The interviews yielded more balanced discussions on the negative and positive aspects of research experiences. The immersive summer program that included cohort elements and intentional scaffolding generally provided different experiences compared to more individualized research during the academic year. The interview findings point to the benefits of rich research methods and raise concerns about placing too much weight on assessments based on responses to agree-disagree type survey items.**

Keywords—undergraduate research; assessment

I. Introduction

A number of beneficial outcomes of undergraduate engagement in research have been identified. These have been grouped into six categories [1]: personal gains, professional gains, thinking and working like a scientist, skills, demonstrating norms of professional practice, and understanding how scientists practice their profession. This large ethnographic study interviewed students, faculty advisors, and administrators at four liberal arts colleges in primarily science-based, immersive summer research programs [2, 3]. The longitudinal perspectives of practicing scientists in physics and chemistry reflecting back on undergraduate research selected the most beneficial outcomes from among ten options as: exposure to genuine research (49%), building confidence to conduct research (16%), and development of lab techniques (15%); the other seven outcome options were selected by 5% or fewer of the respondents [4]. Other large studies have explored NSF-funded Research Experiences for Undergraduates (REU) sites using survey methods with Likert-type items. The study by Hancock [5] determined that the summer research "sites" generally engaged in more activities compared to "supplement" grants to NSF projects that more typically funded only one or two undergraduate students on a project over the academic year. For example, about 40% more of the students in summer sites delivered an oral presentation describing their research/results and 30% more prepared a final written report, compared to the undergraduate students funded by ENG supplements. Other large differences were more training in research ethics and research-related field trips in the summer programs compared to supplements. Survey questions on potential gains in confidence, skills, understanding, and awareness invited responses on a four-point scale (1 = not at all, 2 = somewhat, 3 = a fair amount, 4 = a great deal). Ten of 18 items had an average rating of 3.0 or higher; the lowest average rating was 2.6 (awareness of ethical issues in conducting research). The outcome items were similar among different types of research experiences, with the exception of greater gains in skills by those participating in the summer immersive research programs.

A concern is that survey methods can be subject to positive response bias and self-reported gains by students may be questionable. Richer methods that have been used to study undergraduate research include student reflective essays [6], focus groups [7], and observations [8]. However, few of those studies have focused on engineering students engaging in undergraduate research. Two types of survey response bias may be of particular concern while trying to accurately assess the outcomes of undergraduate research experiences: acquiescence response and social desirability response bias. Acquiescence bias is the tendency to respond to survey items by agreeing. Acquiescence bias has been linked with being agreeable, conforming to authority, and satisficing [9]. A number of factors have been shown to impact acquiescence bias, including gender, age, race/ethnicity, education, conservatism, culture, and country of origin (amount of corruption, collectivism, GDP) [10-12]. Social desirability bias is the tendency to respond to items in a way that seems more

National Science Foundation REU site grants 0552895, 1004122, and 1263385 and EEC grant 1158863.

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

likely to make a positive impression on others. It has been stated that social desirability bias will be minimized when a person feels that they have more privacy and anonymity [13]. Anonymity is less likely to be associated with program assessment when compared with large studies, since summer REU programs typically have an average of 11 students [5]. Despite reassurances that their responses are anonymous, the students may be concerned that negative feedback would find its way to their faculty advisor, which might impact future interactions such as getting a reference letter.

Given this background, the overarching question explored in this research was: Are differences found in the outcomes of undergraduate research engagement based on how the information is gathered? The research results could help inform better ways to assess the outcomes from undergraduate research experiences.

II. METHODS

This research compared two sources of information to determine the impacts of undergraduate research experience for engineering students. These sources were survey methods and interviews not directly exploring research experience outcomes. The comparative data sets explored were convenience samples. The studies were conducted conforming with informed consent protocols approved by the relevant institutional review board for human subjects research.

A. *Surveys to Assess REU site*

The survey findings originated from 8-years assessing a ten-week summer NSF-funded REU site at a single institution that focused on environmental engineering (2006-08, 2010-2012, 2014-2015; n=82). The cohort included 71% females and 56% white non-Hispanics (other race/ethnicities were 28% Hispanic, some Black, Native American, and others). Students self-rated their level of research experience prior to the REU as 39% nothing significant outside of class-based laboratories, 29% moderate (some hourly research but not an independent project), and 32% extensive (a semester or summer program). The REU experience included a seminar series that covered elements including conducting literature searches of scholarly research, responsible conduct of research, and tours of federal research laboratories. Cohort building was explicit, with a number of group activities in both formal and social settings. Students were required to write a research proposal, and at the end of the summer a concluding written report and either an oral or poster presentation.

The REU data set included Likert-type responses to pre and post survey questions and self-reported gains on the end-of-summer survey [14]. The surveys were administered as paper forms and completed on the first day of the program after orientation, and the last day of the program in concluding activities. The pre survey included a series of items that asked students to rate their level of knowledge of a series of topics (from 0 = none to 4 = excellent) and ability in a number of skill areas that primarily mapped to the ABET A to K outcomes (using a scale of 0 to 4); these same items were repeated on the post survey. Starting in 2010, a new series of items were added to the pre-survey that asked students to rate their confidence to perform 13 research-related tasks on a scale of 0 (no confidence) to 100 (extremely confident); these items were repeated on the post survey. The related-samples Wilcoxon signed rank test was used to statistically compare the paired data, using IBM SPSS version 24.

The REU surveys at the end of the summer began with open-ended response questions. Placing these questions before the Likert items was intentional in order to minimize bias to student responses. Responses to two questions are explored in this research: "What was the most valuable aspect of your summer REU experience for you?" and "If you participated in research prior to this summer, how did the CU REU research experience compare to your other research experiences?". Emergent themes were determined by the first author. This was used to develop a code-book. All of the responses were then evaluated using the code book. The open-ended responses from the 2010 surveys were unavailable for examination.

The post-survey also included a number of items that asked students to "rate how much you agree/disagree with the following statements on potential outcomes form your involvement in your summer REU research experience" using a scale of 1 = strongly disagree/was not a benefit for me, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree/strong benefit. There were over 80 statements of this type on the 2006-2012 surveys. Given the large apparent acquiescence response bias noted, the 2014 and 2015 surveys greatly reduced the number of items and switched to a scale of 0 (none), 1 = a minimal amount, 2 = a fair amount, 3 = a lot, 4 = substantially. These survey responses were recoded to a 1 to 5 scale to allow more direct comparison to the earlier survey response data.

B. *Interviews on various undergraduate research experiences*

The student interviews were part of a longitudinal study on social responsibility; 34 students initially majoring in engineering at four different institutions were interviewed midway through their second semester in college with follow-up interviews in each of the subsequent three years. Among the twenty students who persisted in engineering and the research study, nine had a research experience. Some students had multiple research experiences over time. None were participants in NSF-funded REU sites, but some were immersive summer experiences. Given the longitudinal nature of the study, changes in students' descriptions of their research could be determined. Audio recordings of the interviews were transcribed verbatim, and assigned pseudonyms. Transcripts were coded using a combination of a priori codes expected from published literature and grounded codes that emerged from the interviews.

III. RESULTS AND DISCUSSION

A. *Quantitative Data from REU Pre-/Post- Surveys*

Students rated their level of knowledge of 16 topics related to the REU, including general research related elements (e.g. experimental design, responsible conduct of research) and specific environmental engineering related topics (e.g. sustainability). On the pre-survey, students on average rated their knowledge of research and graduate student funding the lowest (average 1.3/4); the specialty topics acid mine drainage (average 1.0/4) and numerical simulation (average 0.9/4) were also rated low. The global need for environmental engineering was rated the highest on the pre-survey (average 2.9/4). The most "saturated" item on the pre-survey was lab safety, with 24% of the students rating this at the top of the scale with a 4. Of the 16 knowledge outcomes, 15 showed statistically significant increases from the pre to the post survey in the related samples Wilcoxon signed rank test. This included 60-70% of the students with increases in knowledge of: experimental design, research/graduate student funding, research methods & data QA/QC, contemporary issues pertaining to environmental engineering, and acid mine drainage. There were 50-60% of the students who increased in their knowledge of: responsible conduct of research, environmental justice, field sampling, impact of engineering solutions in a global and societal context, global need for environmental engineering, air pollution, and numerical simulation. Only 40-50% of the students increased in ratings of their knowledge of: environmental ethics, environmental policy, lab safety, and sustainability (but sustainability gains not statistically significant, p=0.057).

Among ten ABET A-K outcomes, eight showed statistically significant increases in students' self-ratings of their skills between the pre to post survey (2006-2012) in the related samples Wilcoxon signed rank test. Among these, 60-70% of the students increased in their self-rated ability to: use the techniques, skills, and modern engineering tools necessary for engineering practice; and design and conduct experiments. These are logical outcomes given the research environment of the REU. There were 40-60% of the students who increased their ability to: design a system, component, or process to meet desired needs; identify, formulate, and solve engineering problems; analyze and interpret data; engage in lifelong learning; apply knowledge of mathematics, science, and engineering; 33% increased in their self-rated ability to function on multidisciplinary teams. The most saturation on the pre-survey was found in students' self-rated ability to engage in lifelong learning, with 47% rating this at the top of the scale with a 4. Changes were not significant in the cohort overall in their self-rated written communication ability (pre-survey average 3.2, 29% increased) or oral communication ability (pre-survey average 2.9, 26% increased).

There were not significant differences among the cohort as a whole between the pre- and post-surveys on their self-rated likelihood that they would pursue a Master of Science (MS) or Doctor of Philosophy (PhD) degree. For MS, the average pre-survey rating was 3.4/4 (3=good, 4=excellent); 19 students increased in likelihood, 14 decreased, and 29 remained the same. For PhD likelihood the average pre-survey rating was 2.3/4 (2=fair); 20 students increased, 26 decreased, and 26 remained the same.

In a smaller number of years (2010-2012, 2014-2015), students rated their confidence in 13 different research-related skills using a scale of 0 (no confidence) to 100 (extremely confident); results are summarized in Table I, sorted from items with the largest average increase to the smallest increase.

This larger range of response options had more fidelity than the five-point Likert-style scale used in the previous items. The majority of the respondents still used only the upper part of the scale, with 76% to 98% of the students rating each of the 13 items on the pre-survey at 50 or higher; thus, a positive response bias may still be at work. However, some students rated themselves very low in some areas, with minimum responses of 0 to 10 on five of the survey items. The survey at the beginning of the summer showed a very wide range in these self-ratings, not surprising given that the previous amount of research among the students ranged from none outside of courses, up to students who had previously completed full-time research over the summer. Students on average had the greatest incoming confidence in their ability to work effectively with others on a team (average 91; 39% of the students rated this at 100) and the lowest confidence in their ability to design and implement the best measurement approach to study an aspect of environmental engineering (average 63). By the end of the summer, the responses generally increased and converged. Confidence in ability to learn lab/field techniques did not increase among the cohort; this may have been due to the double-barreled nature of the question. While the majority of the REU students performed lab work and learned lab techniques, a few engaged in field work; some worked entirely on numerical modeling and did neither lab nor field research. In terms of teamwork, while some students worked with large research groups (a faculty mentor, post-doc, one or more PhD and Master's level graduate students, and/or other undergraduates), others worked with a much smaller team (most typical was the faculty mentor, one graduate student, and the REU student), and a few worked only with their faculty mentor. The results also appear to show that some students may have better realized the true complexity of research after the immersive REU experience; confidence decreased for 11% to 30% of the students on various research-related tasks (among 47 sets of paired data). This is not particularly unexpected given that some students rated themselves at 100 on the pre-survey.

TABLE I. REU STUDENT CONFIDENCE IN RESEARCH-RELATED ITEMS ON THE PRE- AND POST-SURVEYS

Survey Item	Average Increase in Mean	Pre Mean ± Stdev (range) n=54	Post Mean ± Stdev (range) n=48	% increased	% decreased
Design and implement the best measurement approach possible for your study of some aspect of environmental engineering	17**	63 ± 20 (20-100)	80 ± 11 (50-98)	79	11
Design and implement the best sampling strategy possible for your study of environmental engineering	15**	65 ± 20 (20-100)	80 ± 12 (50-100)	74	17
Use various technological advances effectively in carrying out research	14**	68 ± 23 (10-100)	82 ± 12 (50-100)	74	15
Review a particular area of environmental engineering or science theory and research, and write a balanced and comprehensive literature review	13**	70 ± 23 (10-100)	83 ± 13 (50-100)	64	26
Choose a research design that will answer a set of research questions and/or test a set of hypotheses	13**	68 ± 22 (10-100)	81 ± 11 (50-100)	66	26
Design and implement the best data analysis strategy possible for your study	11**	72 ± 21 (20-100)	82 ± 10 (60-100)	60	28
Formulate a clear research question or testable hypothesis	11*	74 ± 23 (0-100)	85 ± 12 (60-100)	64	26
Do effective electronic database searching of the scholarly literature	10*	75 ± 23 (10-100)	85 ± 13 (40-100)	57	23
Understand the ethical implications of my research	8**	81 ± 15 (40-100)	90 ± 9 (60-100)	60	19
Interpret data by relating results to the original hypothesis	8*	79 ± 17 (30-100)	88 ± 9 (70-100)	62	23
Effectively present your study and its implications	6**	85 ± 15 (40-100)	91 ± 8 (60-100)	57	15
Work effectively with others on a team	4*	91 ± 13 (40-100)	95 ± 7 (60-100)	40	17
Learn lab/field techniques (e.g. instrumentation, measurement)	2	87 ± 15 (40-100)	89 ± 12 (50-100)	43	30

Related-samples Wilcoxon signed rank test, **sig. ≤ 0.001, * sig. ≤ .05

B. Quantitative REU Post Survey Gains

On the post survey, the REU students were asked to respond to items related to the potential broader outcomes of research experiences, such as increased confidence, general knowledge, or mentoring outcomes (83 items) and items on specific environmental engineering topics (n=11, such as fluid mechanics, remediation, drinking water, etc.). Some students left some survey items blank. In the past, these responses were merely missing from the average. However, upon further consideration of the survey responses, it appears that the lack of response may have been because the student did not consider the item to be an outcome from their experience, but for whatever reason felt uncomfortable identifying it as-such. An example is an item that read "My research experience helped me to establish a mentoring relationship with professional engineers." The vast majority of the research experiences did not include interfaces with professional engineers; however, many students just left this item blank rather than entering what would be an appropriate and expected response of 1 (strongly disagree/was not a benefit for me). Therefore, in the current analysis the missing responses are included as "disagree" (2) responses.

These were the series of questions where acquiescence response bias and/or socially desirable response bias were most widely suspected. The first aspect of poor survey design was the fact that from 2006-2012 this portion of the survey included 83 items, among which all but one (item 78) were positive statements. This potentially overwhelming series of questions may have led some students to "zone out" and not carefully consider each individual item. This idea is supported in 12 cases where students' responses to the statement "research experience clarified that a research career is not what you want to pursue" (item 78) seemed to contradict their answer to the statement "interest in a research-related career" (item 72). Across these 83 general items, 60 items had average responses of 4 (agree) or higher (where the maximum was 5 = strongly agree) and another 16 items had average student responses of 3.8-4.0. This seems to be a strong positive bias in the responses; we certainly don't expect all students to experience all of the potential outcomes. As another way to explore the data, 11 of 60 students (18%) rated 90% or more of the 83 items at 4 or 5; another 21 students (35%) rated 80-90% of the 83 items at a 4 or 5.

Dissatisfaction with this portion of the survey led to a reduction from 94 to 26 items on the 2014 and 2015 REU post surveys. In addition, the response scale was changed from 1 to 5 disagree/agree to 0 to 4 for 20 items prefixed by the statement "my research experience increased my…", ranging from 0= none, 1= a minimal amount, 2 = a fair amount, 3 = a lot, 4= substantially. Six items were changed to a 0 to 4 Likert-type scale from 0 = strongly disagree to 4 = strongly agree.

These changes in 2014-2015 perhaps resulted in more realistic responses. Table II reports the average for selected survey items after conversion to the 1 to 5 scale for comparison with the older REU survey data. While the averages did not show large differences, there was generally a smaller percentage of agree responses (the upper end of the scale). For example, in the old survey 98% of the REU students agreed/strongly agreed that the experience increased their confidence; on the revised survey, only 74% of the students agreed with this statement, 21% were neutral, and 5% disagreed.

TABLE II. REU STUDENT POST-SURVEY AGREEMENT & INTERVIEWEE OUTCOMES

Item	REU 2006-12 (n=60)			REU 2014-15 (n=19)^			Student inter-view Denise
	Mean	*% agree (4, 5)*	*% disagree (1, 2, blank)*	*Mean*	*% agree (4, 5)*	*% disagree (1, 2, blank)*	
Confidence	4.4	98	0	3.9	74	5	Y
Confidence in my ability to do research	4.5	97	0	4.3	89	0	Y
Feeling like a scientist	4.3	83	2	4.2	74	0	Y
Critical thinking / problem solving skills, in general	4.3	88	3	4.3	63	0	Y
Patience and perseverance	4.1	77	10	4.2	74	5	Y
Ability to work independently	4.4	88	5	4.4	79	0	Y
Work organization skills	4.1	82	7	3.9	63	0	
Intrinsic interest in learning	4.3	88	0	4.4	89	0	Y
Attention to detail	4.2	87	2	3.9	68	0	Y
Self-esteem	4.2	83	5	4.0	79	11	Y

^converted to 1 to 5 scale from the 0 to 4 scale

Additional survey data supports the idea that the question format may lead to different conclusions. Comparing the likelihood to pursue graduate studies as rated by the students on both the pre and post survey (Table III), only about 30% increased. Many of the students (n=41, 50%) had "saturated" the scale on the pre survey, by selecting the maximum rating of "4" for excellent likelihood to pursue an MS degree. The high initial interest in graduate studies is not particularly surprising, given that applicants wanting to spend an entire summer in a research program are often already fairly committed to the idea of graduate studies and are building their resume for admittance. A notable percentage of the students decreased in their graduate school likelihood ratings the post survey. By comparison, the responses to the Likert-type "level of agreement" items later in the survey showed strong agreement that the REU experience increased the probability that they would pursue graduate degrees (79% agree/strongly agree) and increased their interest in graduate school (83% agree/strongly agree). These are quite different results compared to the likelihood ratings. Multiple items to query a similar idea is good survey protocol; however, the pre-/post- format may be more realistic. The precision of the item that asked students about the likelihood of pursuing graduate degrees (0 = none, 1 = minimal, 2 = fair, 3 = good, 4 = excellent) may have been too limiting. A richer scale of 0 to 100 might be more appropriate.

TABLE III. REU GRADUATE SCHOOL INTEREST

Survey: Item	Increased post vs. pre or agree, %	No change post vs. pre or neutral, %	Decreased post vs. pre or disagree, %
Pre/Post: Likelihood pursue MS	31	47	23
Pre/Post: Likelihood pursue PhD	28	36	36
Post Q72/12: increased the probability that I will go on to graduate school / higher degree	79	18	3
Post Q82: research clarified/confirmed level of interest in graduate school (2006-12)	90	8	2
Post Q19: research experience increased my level of interest in graduate school (2014-2015, n=18)	83	11	6

C. *Qualitative REU Survey Responses*

Open-ended questions are less likely to be subject to various types of response bias. One of the open-ended questions asked students, "What was the most valuable aspect of your summer REU experience?" Themes that emerged from the 61 responses are shown in Table IV, with a number of students citing interpersonal elements, knowledge/skills, and/or personal attributes.

TABLE IV. MOST VALUABLE ASPECTS OF REU THEMES

Theme	n	Example Quote
Cohort / networking	16	Meeting new people
Mentoring from faculty and graduate students	13	Making relationships with faculty and working with them
Understanding research	13	Getting a feeling for a very broad range of issues that go into a research project, especially one conducted in a lab
Lab skills	10	Being able to work in the lab and learn methods of using instrumentation
Understanding graduate school	9	Learning more about grad school was very helpful
Communication skills	9	Technical writing practice
Knowledge and understanding of environmental engineering	9	Just seeing what work in the groundwater field is like
Data analysis skills	5	Learning to analyze data
Independence	5	Having the freedom and latitude to think for myself. It forced me to step up and become a better researcher
Confidence	3	This experience gave me confidence in what I can do and in how research works. Now it's not so intimidating

For the question "If you participated in research prior to this summer, how did the this REU research experience compare to your other research experiences?" there were 34 responses, excluding those that were blank or wrote in "N/A" or have not done previous research. These results primarily represent contrasts with academic year research experiences, although some had previously participated in other REU sites. Among the responses the common themes were:

- better / enjoyed more / great (n=9)
- more independence / freedom / autonomy (n=8)
- more in depth / demanding / intense (n=6)
- better motivation, sense research worthwhile (n=4)
- more focused (n=3)
- learned more (n=3)
- more goal oriented / research quality / output (n=3)
- similar (n=3)
- higher level of faculty involvement (n=3)
- mentor interactions (n=2; uncertain if faculty mentors or other mentors)
- less organized (n=2)
- more structured (n=2)
- cohort experience with other peers doing research (n=1)

A few example quotes that illustrate the themes above are:

"I enjoyed the relationship with my mentor more than my previous research. Also, that there were more worthwhile research and demands."

"This experience was more valuable because I had to figure out a lot of things for myself instead of being told exactly what to do. I think I'm more prepared for graduate school now."

"I enjoyed it more - it was good to be around other undergrads doing research, and my faculty mentor was very involved."

"Much more open with how to design my own experiment - I liked the freedom."

Some of the themes from students' comparisons of their research experiences mirror those found when students' discussed the most valuable aspects of the experiences. The immersive experience of the summer program that allowed more autonomy and depth seemed to be key features, as well as the critical role of the research mentors (faculty and/or graduate students).

D. Interview Results

Among the nine students who discussed undergraduate research experiences, five conducted research only in the academic year and four discussed research over the summer at their home institutions (three among this group also conducted academic year research). Among the five with only academic year research experiences, the interviews yielded more balanced discussions on the negative and positive aspects of undergraduate research experiences. Ashley, a chemical engineering major, went from describing her physical chemistry class lab in her sophomore year as "fun", to talking little about her on-going research during her junior year, to saying in her fourth year "I learned that I don't like academic research or being in a lab. It's not for me." A mechanical engineering student, Madison, was fitting undergraduate research into a busy schedule. In her sophomore year interview she stated, "I work a lot up here. I work in a coffee shop on campus like 13 hours a week and then water polo and broom ball and I'm doing some research with a professor here…." She was asked to describe the research, and then was asked if she found it interesting. Madison replied, "Yeah, I enjoy it. I don't know, I haven't actually been on the project yet so it will be interesting to see what that's like…" Thus, her involvement to date had been planning the experiments. When asked if the experience was different from her technical engineering classes she replied, "Yes different. And kind of similar." Interestingly, Madison did not mention her research experience in the subsequent interviews in her junior and senior year. So it is unclear that the experience had a lasting impact. Tucker, a civil engineering student, discussed conducting research over the previous two academic years during his year 4 interview. Tucker indicated that he enjoyed his research, found it fascinating and pretty satisfying, but "I don't love it, honestly." He described gaining an understanding of the research process, as well as critical thinking/problem-solving skills. Rachel, a computer science major, described the research during her sophomore year with robots as building her resume, while in her junior year discussed elements of working with different kinds of engineering on her research team. In her year 4 interview she did not discuss research, but did have plans to attend graduate school abroad.

There were hints that immersive summer programs that include cohort elements and intentional scaffolding provide different experiences and outcomes as compared to more individualized research during the academic year. In the fourth-year interview, mechanical engineering student Derek stated, "There were a bunch of other students working on different projects through the whole summer and it was really cool to go around and say 'hey, how's your project going?' … it was really neat having that culture and community." The rest of his research discussion tended to focus on specific technical details and the skills that he had learned.

Denise, a mechanical engineering student, had been talking about research since the interview in year 2, when she discussed applying to REU sites for the summer. In year three, she indicated that she had written a research proposal with a professor to work on a project with biomedical sensors, stating, "So hopefully I'll be able to work with him over the summer, and the main reason that I'm pursuing some research is that I am about 95% sure that I want to go to grad school when I graduate. So I'm kind of looking to do a Masters and I want to get some research under my belt before I apply to grad school." Denise described her research experiences in-depth during the fourth-year interview. She described her summer research, which she continued into the following academic year. She had been admitted to the MS program at her undergraduate university. Denise spoke at length about her research experience without significant prompting. When the REU survey items were considered in coding her response, a large number of those potential research experience outcomes were identified. For example, among the 25 "gains" outcomes on the revised 2014-2015 survey, 23 were evident in Denise's interview (a number of these are summarized in Table II). A quote illustrates many of these themes:

> "So it requires a lot of patience and it can take a while to go through, but for the most part it is been a very valuable experience learning about all of these different things in a short period of time and also personally feeling more independent and feeling like I can do this kind of work and that's really important feeling like I'm part of the team. And that's very rewarding because the group is pretty small and getting to know everyone pretty well has been very worthwhile for me making some long-lasting relationships. So for me, the biggest things have been the everyday personal interactions besides the research itself which has gone really well."

Denise also discussed lab safety, responsible conduct of research, learning lab techniques, and research/graduate student funding.

The interview findings point to the benefits of rich research methods, as well as differences between immersive participation in research over the summer versus research during the academic year.

IV. Conclusions and Implications

It is clear from the results that not all research experiences are the same. This was also found in a recent review of STEM undergraduate research experiences [15]. The REU participants noted some of these differences in their qualitative responses to the end-of-summer survey, and differences were also evident among the interview responses. A number of academic-year research positions seem to yield fewer benefits as compared to immersive experiences over the summer. The REU experience has the benefit of the faculty PI and/or co-PIs who may serve as "global mentors" to the research experience, organizing structured activities and deliverables that frame the overall research experience; these individuals can help facilitate a good experience for all of the students in the cohort, perhaps being more approachable for students to share difficulties and frustrations in comparison to the faculty mentor for their personal research project. Students in the immersive experiences seem to take more ownership for their projects, engaging thoughtfully at the start via the literature search and proposal process, through data analysis, and writing up their findings in a report and making a public presentation (either oral and/or poster) on their findings. More often, the academic year experiences have undergraduate students dropping in to an on-going project, and "contributing" without a full understanding of the big picture of the project or feeling independently responsible for achieving particular outcomes. At the REU site, the cohort of about ten undergraduate students each summer frequently both live and work together, so they have lots of opportunity to discuss their projects informally with each other and learn from the broader suite of projects. This contrasts with a typically narrower view that a student will have when working during the academic year with a team on a particular project. In academic year experiences, each research project typically includes only a single undergraduate, so they lack a cohort experience with other undergraduates. In addition, the summer REU experience is the primary focus for the undergraduate participants; the academic year experiences typically require that students juggle a full-time class load with their research engagement. This immersion of about 40 hours per week differs from typically 10 hour per week research involvement during the academic year.

Faculty may wish to advise students with potential interests in research to try a smaller academic year commitment, before pursuing an immersive summer experience. Some students are able to quickly realize that research is not a good fit to their interests, and that revelation is better for all parties (students and their range of research mentors) from a smaller investment. REU sites can feel fairly confident of helping students to realize significant benefits from a first immersive summer research experience, and resources should be devoted to these "novices". Undergraduates who have already enjoyed a summer REU site experience should pursue other types of research and/or internship experiences rather than another REU, which is likely to be quite similar. After an REU experience, these undergraduates can likely contribute to research at the level equal to or exceeding an incoming master's student without an independent research experience. Therefore, standard research funding to support these experienced researchers seems warranted.

Mentors of research experiences may want to incorporate reflective essays [6] or structured cognitive reflection journals [16] as a required element. The process of reflection itself can stimulate metacognition, which results in improved learning outcomes [17]. Therefore, reflective essays can serve a dual role as both a method to enhance student learning and as an assessment tool. A best practice may be to require initial, midpoint, and final reflections or weekly reflective journaling in place of pre/post Likert-type surveys. To ensure more honest reflection, the direct research mentors should not read the journals. Rather an individual involved in program assessment, the REU site PI, or a graduate student not serving as a primary mentor could evaluate the content of the reflections / journals. The feedback from the undergraduates could be viewed as formative assessment for the research program and enable mid-course correction to ensure the best learning experiences for the undergraduate student participants.

Individuals are cautioned that the types of surveys typically used to assess summer undergraduate research programs may be prone to biased results that may over-state the beneficial impacts on students. Surveys with a number of "agree-disagree" items are likely to be prone to acquiescence and/or social desirability response bias. Students likely feel pressure to rate the experience positively, given the small cohorts where anonymity is difficult and perhaps particularly if they hope for recommendation letters from their mentors. If the students generally had a good experience, they may bring this positive feeling to the survey as a whole, without critically analyzing their personal outcomes. Programs may generally prefer this wealth of positive feedback, as it is advantageous to report this data to funding agencies.

For those who wish to gather more accurate assessment of their program and/or individual mentoring, there are a number of recommendations. Items where responses can be compared on pre and post surveys can temper issues with positive response bias. The self-ratings on a wide scale, such as 0 to 100, allow more resolution as compared to typical 5-point scales. The literature suggests that a balance of positively and negatively worded items could help reduce acquiescence bias, or presenting response options from both strongly disagree to strongly agree and vice versa [18]. Also, a smaller number of items may yield improved results as students will feel less pressure to move rapidly through numerous items versus carefully moving through the four-stage cognitive process to survey response that has been proposed [19]. Overall, the most accurate results may be gathered from open-ended response items. This is less likely to introject the bias from the hoped-for outcomes. However, analysis of these responses is more complex. Although various methods to identify themes and code responses are well established [20], students' language and tone are sometimes unclear. Elements grouped into similar themes may also have important, nuanced differences. A richer reflective essay might provide even better information than short response items. While data analysis is more difficult from reflections, it is expected that the quality of information will be more accurate and meaningful. Trading information quality for quantity is likely to be worth the effort for undergraduate research programs and/or individual mentors who wish to have a more accurate understanding of their impact on students.

REFERENCES

[1] A-B. Hunter, T.J. Weston, S.L. Laursen, and H. Thiry. "URSSA: Evaluating student gains from undergraduate research in the sciences," Council on Undergraduate Research Quarterly, vol. 29, no. 3, pp. 15-19, 2009.

[2] A-B.Hunter, S.L. Laursen, and E. Seymour. "Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development," Sci. Ed., vol. 91, pp. 36-74, 2007.

[3] E. Seymour, A-B. Hunter, S.L. Laursen, and T. Deantoni. "Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study." Science Education, vol. 88, no. 4, pp. 493-534, 2004.

[4] J.A.Harsh, A.V. Maltese, Rand .H. Tai. "Undergraduate research experiences from a longitudinal perspective," Journal of College Science Teaching, vol. 41, no. 1, pp. 84-91, 2011.

[5] M.P. Hancock and S.H. Russell. Research Experiences for Undergraduates (REU) in the Directorate for Engineering (ENG): 2003-2006 Participant Survey. Report Prepared for the National Science Foundation. 132 pp. 2008.

[6] A. Hosein and N. Rao. "Studemts' reflective essays as insights into student centered-pedagogies within the undergraduate research methods curriculum," Teaching in Higher Education, vol. 22, no. 1, pp. 109-125, 2017.

[7] N.G. Holmes and C.E. Wieman. "Examining and contrasting the cognitive activities engaged in undergraduate research experiences and lab courses," Physical Review Physics Education Research, vol. 12, paper 020103, pp. 1-11, 2016.

[8] J.J. Thompson, E. Conaway and E.L. Dolan. "Undergraduate students' development of social, cultural, and human capital in a networked research experience," Cult. Stud. Of Sci. Educ., vol. 11, pp. 959-990, 2016.

[9] O. Kuru and J. Pasek. "Improving social media measurement in surveys: avoiding acquiescence bias in Facebook research," Computers in Human Behavior, vol. 57, pp. 82-92, 2016.

[10] G. Marin, R.J. Gamba, and B.V. Marin. "Extreme response style and acquiescence among Hispanics: the role of acculturation and education," Journal of Cross-Cultural Psychology, vol. 23, no. 4, pp. 498-509, 1992.

[11] B. Rammstedt, D. Danner, and M. Bosnjak. "Acquiescence response styles: a multilevel model explaining individual-level and country-level differences," Personality and Individual Differences, vol. 107, pp. 190-194, 2017.

[12] J. Welkenhuysen-Gybels, J. Billiet, and B. Cambre. "Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items," Journal of Cross-Cultural Psychology, vol. 34, no. 6, pp. 702-722, 2003.

[13] D. Dodou and J.C.F. de Winter. "Social desirability is the same in offline, online, and paper surveys: A meta-analysis," Computers in Human Behavior, vol. 36, pp. 487-495, 2014.

[14] A.R. Bielefeldt, A.R. 2012. "Student learning outcomes from an environmental engineering summer research program," American Society for Engineering Education (ASEE) Annual Conference and Exposition Proceedings. San Antonio, TX. June 10-13, 2012. 21 pp. https://peer.asee.org/21943

[15] J. Gentile, K. Brenner, and A. Stephens, Eds. Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities. Prepublication Copy. National Academies Press, Washington DC. 2017.

[16] S-K. Lee and Mina Choi. "A study on effects of well-structured cognitive reflection journal on metacognition and learning achievement," Journal of Engineering Education Research, vol. 11, no. 1, pp. 5-14, 2008.

[17] M. Kaplan, N. Silver, D. Lavaque-Manty and D. Meizlish. Using reflection and metacognition to improve student learning: across disciplines, across the academy. Sterling VA: Stylus Publishing, 2013.

[18] J.J. Barnette. "Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems," Educational and Psychological Measurement, vol. 60, no. 3, pp. 361-370, 2000.

[19] M. Yorke. "'Student experience'surveys: some methodological considerations and an empirical investigation," Assessment & Evaluation in Higher Education, vol. 34, no. 6, pp. 721-739, 2009.

[20] J. W. Creswell, Qualitative Inquiry & Reserach Design: Choosing Among Five Approaches, Thousand Oaks, CA: Sage Publications, 2007.