

Evaluating a Scalable Program for Undergraduate CS Research

Christine Alvarado
UC San Diego
La Jolla, CA
cjalvarado@eng.ucsd.edu

Sergio Villazon
UC San Diego
La Jolla, CA
svillazo@ucsd.edu

Burcin Tamer
Computing Research Association
Washington, DC
burcin@cra.org

ABSTRACT

Undergraduate research experiences have been shown to have many positive effects on undergraduates including increased confidence, sense of belonging and retention. However, many previous studies of undergraduate research experiences have focused on advanced undergraduate (juniors and seniors) in one-on-one research experiences with a faculty mentor. Less is known about the effects of early undergraduate research, particularly via opportunities that scale beyond one-on-one faculty-student relationships to encompass large numbers of early undergraduates. The research question addressed in this work is whether a more scalable group-based research model aimed at early undergraduates from groups underrepresented in computing would show the same kinds of benefits for participants as more personalized one-on-one programs aimed at more advanced students. We evaluated a group-based early research program in the computer science department of a large public university. Through survey data and direct measurements of performance and retention several years after students had completed the program, we found that students who participated in this program have higher overall GPAs, more confidence, and more interest in research compared to several different control groups. Our design also allowed us to examine the considerable impact that selection bias can have on the evaluation of research programs. This work both validates the scalable structure of this research program and provides a richer perspective on the benefits of early undergraduate research in CS.

CCS CONCEPTS

• **General and reference** → **Evaluation**; • **Social and professional topics** → *Computer science education*;

ACM Reference Format:

Christine Alvarado, Sergio Villazon, and Burcin Tamer. 2019. Evaluating a Scalable Program for Undergraduate CS Research. In *International Computing Education Research Conference (ICER '19)*, August 12–14, 2019, Toronto, ON, Canada. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3291279.3339406>

1 INTRODUCTION

The benefits of undergraduate research are well-documented and include increased retention, increased sense of belonging in the

major, and increased performance [2, 6, 7, 20, 21]. These positive effects may be due to an increased self-efficacy or an increased sense of belonging, both of which have been shown to be important in students' success [1, 8, 9, 27, 28], and this may be one reason why research has been shown to have particularly strong benefits for students from groups underrepresented in STEM [16, 25].

Recent work on undergraduate research experience in computer science (CS) specifically has shown that these benefits are not limited to upper-class CS students who engage in research, but that similar benefits are present even for early-career undergraduates [23]. Yet it can be difficult for first and second-year undergraduates to find research opportunities. The few opportunities that do exist usually go to very advanced early undergraduates who have significant pre-college CS experience (which fewer students from underrepresented groups have [5, 13, 14]), or require a level of support that most mentors are unable to provide.

The Early Research Scholars Program (ERSP) was created in 2014 in the Computer Science and Engineering (CSE) department at UC San Diego (UCSD) to provide early research opportunities to a larger and more diverse group of CS students. ERSP annually provides over 40 mostly second-year undergraduates, the majority of whom are from underrepresented groups, with an academic-year long research apprenticeship in CS.

In order to scale, ERSP modified some of the components present in standard Research Experiences for Undergraduates (REUs), and these modifications potentially make ERSP less beneficial to its participants. First, students in ERSP work in groups, rather than primarily in individual research. This group-based structure could affect students' feelings of personal accomplishment in the program or introduce complicated group dynamics issues that negatively affects their experience. Second, the mentoring relationship is split between multiple advisors. Students might not develop the close sense of support that a more personalized one-on-one advising relationship could give them. Third, being so early in their studies, students might be more overwhelmed than empowered when thrown into an advanced research project.

The central goal of our research was to determine whether the ERSP program has similar positive benefits for participants as other REU programs. This paper presents a multi-faceted evaluation of the outcomes for ERSP participants from the program's first four years. We analyzed data from two different surveys administered at various points after students participated in ERSP as well as grade and retention data from 1-3 years after ERSP. We found that ERSP seems to have many of the same benefits of traditional, more personalized REUs (and even some benefits over traditional REUs), including high retention in the CS major, an increased sense of confidence and belonging in CS, and higher academic performance for participants over non-participants. The design of our study also allowed us to measure the extent to which these benefits might be

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICER '19, August 12–14, 2019, Toronto, ON, Canada

© 2019 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-6185-9/19/08...\$15.00

<https://doi.org/10.1145/3291279.3339406>

due to selection bias; we found that selection bias does account for some of the differences between participants and non-participants, but our analysis suggests there might be benefits of ERSP even when completely controlling for selection bias.

2 RELATED WORK

Research on the benefits of undergraduate research is not new, yet it is surprisingly diverse in its nature. A recent report by the National Academies of Science, Engineering and Medicine summarized the work that has been done to evaluate undergraduate research programs and found a huge array of evaluation techniques and instruments [15]. That report provides a comprehensive and detailed review of the literature; here we summarize the previous work most relevant to our current study.

By far the most common method used to measure the value of undergraduate research is via surveys or interviews. Both the National Academies report and another recent survey paper of 60 empirical studies of undergraduate research [10] note that most studies use only students' own perceptions of their learning gains. While these indirect measures have their limitations, they at least provide a window into what students think they are getting out of a research experience, and allow researchers to both measure change in students' perceptions and compare the relative perceptions of different groups of students.

The Survey of Undergraduate Research Experiences (SURE) is an instrument designed to measure different aspects of students' experiences [11]. In the first study using SURE, Lopatto found that REU participants from 41 institutions reported gains in many areas including understanding research and techniques, and confidence. He also found a correlation between higher reported benefits from REUs and an intent to continue in a research-oriented career. In a later study Lopatto found that students from underrepresented groups reported higher learning gains than those from majority groups, while women and men reported similar gains [12].

Other survey-based studies compare REU students to their non-REU peers. A large scale survey-based longitudinal study by Russell et al. analyzed data from 3,300 undergraduate STEM bachelors degree recipients. They found that REU experiences increased students' reported understanding, confidence, and interest in a research career compared to students who had not participated in research [20]. In a similar but smaller study, Bauer and Bennett found similar effects [3].

Focusing on research in computing, the Computing Research Association's Center for Evaluating the Research Pipeline (CERP) has developed a national survey instrument and collected data from thousands of CS students at dozens of schools. Stout, Tamer, et al. have analyzed the results from this survey and found that research experiences in computing, even in a students' first or second year of an undergraduate program, can lead to students being more interested in pursuing a research-focused career compared to those who have not had research experience [23, 25, 26].

One of the weaknesses of student surveys is that students may not have an accurate perspective of how much they have truly learned or improved their skills. To address this issue, Singer and Zimmerman designed a multi-stage survey for both student participants and faculty mentors [22]. Comparing student and mentor

responses, they found that students tend to over-estimate their abilities compared to mentors' assessments, but that mentors' nevertheless saw growth in students' abilities.

Beyond surveys, interview-based studies have found that students report many of the same benefits cited in the literature including personal and professional gains and increased plans for grad school [21]. An interview study of CS REU participants by Barker revealed important insights about the structure of the research group and the relationship between the student and the advisor that could make the research experience either positive or negative for the female participants [2].

Some studies more directly measure effects of research through student retention and performance. Seminal work by Gregerman et al. evaluated the impact of the long-standing Undergraduate Research Opportunity Program (UROP) at the University of Michigan. They compared retention UROP participants to non-UROP participants and found that UROP participants had significantly higher retention rates, even when compared against a matched control group with similar pre-college academic characteristics [6]. Bowling et al. reported similar high-retention results for an undergrad research program at the University of Northern Kentucky [4].

Finally, several previous studies (including some mentioned above) have focused on the effect that undergraduate research can have on women and students from underrepresented racial and ethnic groups in STEM. Nyame-Mensah modeled the relationship between student learning and self-efficacy gains from research and students' status as an underrepresented racial minority and found that the benefits for underrepresented students were stronger than those for non-underrepresented minority students [16]. Summers and Hrabowski studied the success of the Meyerhoff Scholars Program at the University of Maryland, Baltimore County, a program that includes a research component that is designed to increase African American students' participation in science and engineering [24]. Pender et al. linked participation in summer research to minority students' graduate program attendance [18]. Specifically in CS, Tamer and Stout have shown the relationship between women and minorities' participation in computing research and their interest in a computing research career [25].

While the landscape of research on the outcomes of REU experiences is vast, most of this previous work takes data from a single source (e.g. surveys), and none focuses specifically on group-based research programs for early undergraduates in CS. Our work contributes to the landscape by studying an undergraduate research program with a novel, scalable structure using a variety of both direct and indirect measures to study the impact on participants.

3 ERSP STRUCTURE

ERSP was designed to support a large number of early undergraduates who possess only the "typical" amount of computing knowledge. To be eligible for ERSP, students must be in their second¹ year at UCSD (either as a direct-admit or transfer student) in a CS or CS-related major, and must not have completed any upper division coursework prior to starting the program. The program has grown over its five years; in its steady-state (years three and

¹There are occasional exceptions to the second-year requirement for students who started their CS major late

beyond) approximately 40 students per year are admitted to do research with approximately 10 faculty mentors. These numbers represent about 10% of the students and 20% of the faculty at UCSD.

ERSP runs for a full academic year. In the first term students take an introduction to research course, the outcome of which is a group-based research proposal for a project that they complete during the remainder of the program.

Students in ERSP are admitted centrally, grouped into teams of two to five (usually four), and matched with a research group in the CS department. Throughout the program, ERSP participants are dual-mentored by a faculty or graduate student technical advisor in the research group the team is matched with and a central ERSP mentor and graduate student assistant. Each mentor meets weekly with the student teams. The ERSP central mentor and graduate student provide general mentoring support (e.g., goal setting, communication skill development), while the technical advisor provides guidance specific to the project. This dual-mentoring structure along with the introduction to research course provides the extra support that early undergraduates need while relieving much of this mentoring responsibility from the technical mentors.

ERSP is a competitive program with typical acceptance rates of around 30%. Applicants are scored not on their academic performance directly, but on their motivation and contributions to diversity as reflected in their application essays. In its first four years, students were selected into ERSP using a threshold + lottery system. After reviewers scored participants, those with scores over a certain threshold were named finalists for the program. The ERSP participants were then randomly selected from the set of finalists.²

4 RESEARCH QUESTIONS

ERSP was created with the hypothesis that it would be a sustainable way to engage a large number of early undergraduate students in meaningful research, and that ERSP the experience would result in positive outcomes for participants, similar to those reported in previous studies on research experiences. It is clear that the first part of the hypothesis is supported: each year participants complete meaningful research that they present at an undergraduate research poster session in the spring and sometimes at external conferences.

However, the second part of the hypothesis pertaining to the effects that participation in the program has on students requires more careful study. Our current work was guided by three research questions:

- RQ1: Does participation in ERSP correlate with higher student retention and/or performance?
- RQ2: Do students who participate in ERSP feel more confident and included in the CS community?
- RQ3: Does ERSP influence students' desire to continue with research and/or attend graduate school?

To explore these questions, we performed a multi-faceted study that collected and analyzed data from sources appropriate to each research question.

²In the fifth year the set of finalists was not large enough to support a lottery.

Table 1: Demographics for ERSP participants in our analysis.

Cohort (N)	Female / Male	Minority / Majority
2014-15 (16)	69% / 31%	50% / 50%
2015-16 (28)	75% / 25%	39% / 61%
2014-15 (34)	65% / 35%	26% / 74%
2015-16 (36)	56% / 44%	31% / 69%
Total (114)	65% / 35%	34% / 66%

5 RQ1: RETENTION AND PERFORMANCE

Does participation in ERSP correlate with higher student retention and/or performance?

Summary: We compared the overall GPA and the retention in a CS major to two different matched control groups.

5.1 RQ1 Population and Data

The central population in this study was the ERSP participants from the first four cohorts (academic years 2014-15 through 2017-18). Table 1 shows the number, gender and racial/ethnic minority status composition of the ERSP participants in each cohort. Gender and race/ethnicity were self-reported by ERSP participants. In this study underrepresented minority students included students who identified as Black, Latinx, and/or Native American/Native Pacific Islander while White and Asian students were considered majority students³. We did not have any non-binary students in our analysis so the only gender categories were female and male. We note that the demographics of ERSP participants are very different from the demographics of the CS major at UCSD, which comprises only 18% women and 12% racial/ethnic underrepresented minorities.

We excluded from our analysis students who declined or withdrew from ERSP at any point in the program because they did not get the full ERSP experience. In all, 39 students in the first four cohorts fall into this category: 15 of these declined their admission before the program started, and 24 withdrew from the program before its end. 18 of these 24 students withdrew less than half way through the program, before the research really got underway. They cited a wide variety of reasons with withdrawing including wanting to pursue more challenging research experiences, being too busy with other priorities, or personal issues. We include a discussion of the outcomes for these students when we present results.

We defined two different control populations for our analysis which together give us a more complete understanding of the effects related to ERSP vs. the selection bias of the students who choose to apply to the program. Our first control group comprises the finalists who were not randomly selected to participate in ERSP (see Section 3). We refer to this group as “Finalists”; demographics of this group are given in Table 2.

Although the finalists provide a rigorous comparison group, the group is small, making it difficult to interpret statistical comparisons. To construct a larger control group, we used propensity matching to build a group of 784 students who did not participate

³This is inline with the National Science Foundation (NSF)'s definition of racial/ethnic groups underrepresented in STEM fields. See, <https://nces.nsf.gov/pubs/nsf19304/digest/glossary-and-key-to-acronyms>.

Table 2: Demographics for ERSP Finalists, the first control group in our analysis.

Cohort (N)	Female / Male	Minority / Majority
2014-15 (14)	50% / 50%	50% / 50%
2015-16 (3)	33% / 67%	33% / 67%
2016-17 (16)	63% / 38%	44% / 56%
2017-18 (5)	20% / 80%	0% / 100%
Total (38)	50% / 50%	39% / 61%

in ERSP. We matched the proportion of students in this group to the proportion of ERSP participants on incoming year, gender, major, and minority/majority status. We refer to this second control group as the “Comparison” group. This larger Comparison group not only provides a larger group for statistical comparisons, but also allows us to study how much of any observed differences are due to selection bias even after controlling for several factors.

Following our approved IRB protocol, we obtained the following deidentified data relevant to this study for each student in the groups described above: incoming term (to RU), admitted major, current major, current overall GPA, gender and race. Information about which cohort a student was a participant or finalist in, when applicable, was included as a field in the deidentified data set. Data was collected at the start of summer 2018 so reflects a different amount of time since the program for each cohort (i.e. three years for cohort 1 and immediately after for cohort 4).

5.2 RQ1 Results

5.2.1 CS Major Retention. To study the relationship between ERSP participation and major retention, we compared ERSP participant major retention to major retention of Finalists and Comparison students. For this analysis we included only students who entered UCSD as declared CS or related majors including Math-CS, Cognitive Science with a specialization in computation or human-computer interaction, and Interdisciplinary Computing and the Arts.

We defined major retention as still being listed in CS or a CS-related major at the time of data collection, even if that major was different from the student’s original major. For example, a student who started as a Math-CS major and switched to a CS major would be considered to be retained, while a student who started as a Math-CS major and switched to an Econ major would not. This definition of retention is likely to overestimate retention for all groups: It will include students who no longer planned to get a degree in CS but who had not yet formally switched out, because there is no strong incentive to do so until the student wants to graduate.

We used two-proportion Z-tests to compare ERSP participants to each of the control groups. Table 3 gives the results. Not surprisingly, retention rates are quite high across all groups, for the reasons listed above, as well as because CS at UCSD is a limited major: once a student leaves it is extremely difficult to get back in. Retention rates are similar between ERSP participants and Finalists, but statistically significantly lower for the broader Comparison group.

As mentioned previously, our ERSP group included only students who completed the entire ERSP program. In the current analysis, this excludes 10 students who were admitted to UCSD as CS or

Table 3: Retention rates for students initially admitted as CS or CS-related majors. z-score and p-values are given for ERSP vs. Finalist and ERSP vs. Comparison two-proportion Z-tests.

Group	init. N	Retention	z-score	p-values
ERSP	56	98%	N/A	N/A
Finalist	21	95%	0.73	0.46
Comparison	615	84%	2.90	0.004*

Table 4: Mean GPAs for students in each group in our analysis. p-values and Cohen’s d are given for the comparison between the ERSP participants and each of the control groups.

Group (N)	Mean GPA	p-values	Cohen’s d
ERSP (114)	3.54	N/A	N/A
Finalist (38)	3.40	0.04*	0.41
Comparison (784)	3.26	< 0.001**	0.918

CS-related majors and also admitted to ERSP but did not complete it. 100% of these students remained in CS or a CS-related major.

5.2.2 Academic Performance. We next examined the relationship, if any, between ERSP participation and overall academic performance. To examine this question, we compared the overall GPAs of ERSP participants to each of the control groups using two-tail *t*-tests assuming unequal variance. We used Cohen’s *d* with pooled variance to measure effect sizes. Results are shown in Table 4.

These results show that on average ERSP participants’ overall GPAs are significantly higher than students in the Finalist and Comparison groups. The effect size is small to medium for the Finalist vs. ERSP analysis and large for the Comparison vs. ERSP analysis. While the difference between the Finalist group and the ERSP group is statistically significant, it is only slightly below the 0.05 significance level. So while it suggests that there is a difference, more data is still needed.

For this analysis we again checked the 39 ERSP participants who were admitted to ERSP but were excluded from the analysis because they did not complete the full program. The average GPA for these students was 3.39 which is nearly identical to the Finalists and was not quite statistically significantly different from the ERSP participants ($p=0.08$).

6 RQ 2: SELF-EFFICACY AND BELONGING

Do students who participate in ERSP feel more confident and included in the CS community?

Summary: We used data from a national survey of undergraduates to compare ERSP participants’ self-reported confidence and belonging to that of students from other REU programs and students who had not participated in an REU.

6.1 RQ2 Study population

To answer our second research question, we conducted a comparative examination of ERSP students’ levels of confidence (i.e., self-efficacy) in their ability to succeed in computing and their sense belonging to the field. These analyses used a national sample of

Sense of belonging

- (1) I feel like I belong in computing
- (2) I feel like an outsider in the computing community
- (3) I feel welcomed in the computing community

Self-efficacy: I am confident that I can ...

- (1) find employment in my area of computing interest
- (2) become a leader in the field of computing
- (3) get admitted to a graduate computing program
- (4) complete my undergraduate degree in computing

Figure 1: Survey items used to measure belonging and self-efficacy. All items were measured on a 5-point Likert scale, Strongly disagree (1) to Strongly agree (5)

survey data collected by the Computing Research Association's (CRA) Center for Evaluating the Research Pipeline (CERP) as part of the Data Buddies project. Using this survey data, we were able to select a comparable sample of students to assess ERSP participants' psychological outcomes related to our second research question.

The analyses presented in this section compares ERSP students to two separate comparison groups: (1) students who participated in other REU programs (Other REU) and (2) students who have not participated in any REU programs (No REU). As with the study populations used for our analyses of RQ1 and RQ3, only those ERSP students who completed the program were included in the analyses presented in this section. The two comparison groups were identified using survey questions regarding students' REU participation. Students who participated in an REU that takes place during the academic year and at the students' home institution were included in the Other REU group while students who reported that they had not participated in any REU program were included in the No REU group. Additionally, these two comparison groups included only students attending PhD granting institutions and majoring in a computing field. We used propensity score matching to select a sample of students that are comparable to the ERSP participants in terms of their gender, underrepresented minority status, first generation college student status, and academic class. [17, 19].

ERSP students' self-efficacy and sense of belonging were measured in the fall of the academic year following their participation in ERSP using the relevant survey items from the national survey, which are shown in Figure 1. Data for the comparison groups were taken from the survey conducted at the same time as the ERSP students. The fourth cohort of participants had not been surveyed at time of these analyses, so we included only the first three cohorts of ERSP in our analysis. The group averages were compared using *t*-tests separately for each cohort. The comparisons with $p < .05$ are considered "statistically significant".

6.2 RQ2 Results

6.2.1 Self-Efficacy. In terms of self-efficacy, results for the first two cohorts generally showed no statistically significant differences between ERSP students and either of the comparison groups, with

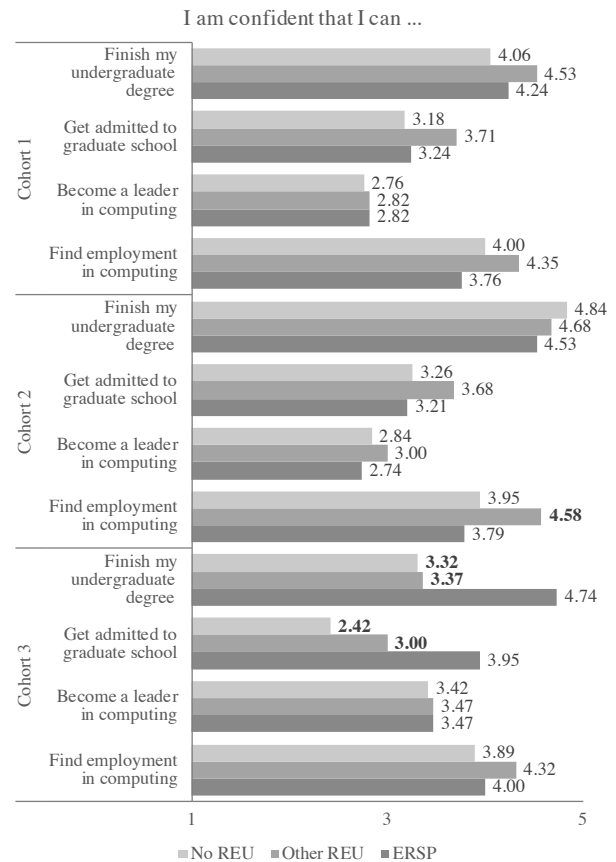


Figure 2: Group averages for the self-efficacy survey items on a 5-point Likert scale, Strongly disagree (1) to Strongly agree (5). Bold font indicates that the mean value of the item for the respective group is statistically significantly different ($p < .05$) from that for ERSP participants.

the exception that ERSP students in Cohort 2 felt less confidence in their ability to find employment compared to Other REU students. However, Cohort 3 participants were more confident that they could complete their undergraduate degree and get admitted to a graduate program than both of the comparison groups. There were no other statistically significant differences in terms of students' self-efficacy. Mean values for each of the survey items measuring self-efficacy are illustrated in Figure 2. The statistically significant comparisons ($p < .05$) are identified with bold font.

6.2.2 Sense of Belonging. Similar to the results on self-efficacy, measures of the first two cohorts' ERSP participants' sense of belonging showed no statistically significant difference between the participants and the two comparison groups, while participants in the third cohort felt that they belonged in computing at a greater rate than both comparison groups. Further, Cohort 3 ERSP students did not feel as much like an outsider in the computing community as Other REU students did. Figure 3 shows the averages for each of the survey items used to measure students' sense of belonging.



Figure 3: Group averages for the sense of belonging survey items on a 5-point Likert scale, Strongly disagree (1) to Strongly agree (5). Bold font indicates that the mean value of the item for the respective group is statistically significantly different ($p < .05$) from that for ERSP participants.

The statistically significant ($p < .05$) comparisons are depicted in bold font.

6.2.3 Underrepresented minorities in computing. As noted in Section 2, several studies reported stronger impacts of REU programs on underrepresented students in computing. In light of this, we examined the ERSP students' self-efficacy and belonging by gender and minority status using the definitions provided in Section 5. Specifically, we compared women in ERSP to the two comparison groups, and minority students in ERSP to the two comparison groups. In a nutshell, in earlier cohorts, women and minority students in the ERSP group had no significant differences than those in the Other REU and No REU groups. For cohort 3, the comparison of results for women who participated in ERSP to women in the two comparison groups matched results for the overall sample reported earlier in this section. Among the cohort 3 students, there was only one minority student who responded to the survey. As a result, analysis of cohort 3 by minority status was not possible.

7 RQ3: CONTINUED RESEARCH

Does ERSP influence students' desire to continue with research and/or attend graduate school?

Summary: We surveyed ERSP participants 1-3 years after their participation to gauge the influence of ERSP on their post-RP research and career choices.

- (1) In what cohort did you participate in ERSP?
- (2) Have you graduated or will you graduate from UCSD by the end of summer 2018?
- (3) After finishing your degree at RU, what did you do, or what do you plan to do?
- (4) After ERSP, did you do any additional undergraduate research outside of a formal course?
- (5) Have you ever applied for a graduate program (in any field)? If so, which one(s)?
- (6) If you have never applied to a graduate program, are you thinking about applying to any graduate program in the next 5-7 years?
- (7) Have you ever been enrolled (or will you be enrolled in the next 6 months) in any graduate program (in any field)? If so, which one(s)?
- (8) If you attended or applied to attend a graduate program, when did you begin or when would you have begun your graduate studies?
- (9) Reflecting on your ERSP experience, how, if at all, do you think your ERSP experience shaped your future career trajectory?
- (10) Do you have any other thoughts about ERSP that you would like to share?

Figure 4: ERSP follow-up survey items

7.1 RQ3 Approach

To study to what degree, if any, ERSP encouraged students to continue pursuing research opportunities and potentially pursue graduate education, we followed up with students in ERSP Cohorts 1-3 in the summer of 2018. We surveyed these students about their post ERSP activities as well as their retrospective views on how ERSP had influenced their career trajectory (if at all). The survey data is both quantitative and qualitative. Here we report descriptive statistics and use the open-ended responses to illustrate impacts of the program on its participants.

Figure 4 gives the specific survey items. Each item except for (9) and (10) included several options for students to select from to give us more information about the students' activities. For example, for question (4) options included "I continued after ERSP with the same professor/project," "I did another research project as an undergraduate at UCSD," "I did another research project at an institution OTHER THAN UCSD while I was still an undergrad," "No, but I am planning to/hope to do additional undergraduate research before I graduate," and "I did not do any additional research as an undergraduate beyond ERSP and I am not planning to." Due to space constraints, options for other questions are not given, but can be inferred from the discussion of the results.

The overall response rate to the survey was 81% with cohort-specific response rates of > 100%, 79% and 71% for Cohorts 1-3, respectively. Cohort 1's > 100% response rate is most likely caused by a student incorrectly reporting their cohort. Unfortunately we have no way of determining which response was in the wrong cohort, so we analyze the data as it was reported.

Table 5: Survey Results. GS = Grad school

Cohort (N)	Addl. res.	GS interest	GS apply	GS attend
2014-15 (17)	76%	88%	29%	12%
2015-16 (22)	41%	73%	23%	5%
2016-17 (24)	58%	88%	30%	17%

7.2 RQ3 Quantitative Results

Table 5 gives the percentage of respondents from each cohort who reported doing additional research, being interested in grad school in a computing field, applying to grad school in a computing field, and having been enrolled in a graduate program in a computing field. The percentage of students who reported doing post-ERSP research is taken directly from survey question (4). The percentage of students who reported being interested in attending grad school (GS interest) is the union of students who reported that they had applied to grad school in question (5) or who were thinking about applying to grad school in question (6). The percentage of students who applied to grad school (GS apply) is taken directly from question (5). Finally, the percentage of students who have been enrolled in grad school (GS attend) is taken from students' answers to question (7).

These results show many trends. A large number of ERSF students went on to do additional research. 15 out of the 36 students who did additional research beyond ERSF continued their ERSF project past the end of the program, either into the summer or into the following academic year(s). 26 of the students who did additional research (perhaps also) did a separate research project not related to their ERSF project. Finally, of the 27 ERSF participants who did not do another undergraduate research project, three had not yet graduated at the time of the survey and hoped to do additional research before they graduated.

A large number of ERSF students (52 out of 63) expressed interest in applying to graduate school in computing in the near to mid future. Most of these students (41) expressed interest in applying (or had already applied) to a Masters' program in computing, but some (12) were (also) interested in pursuing a PhD in computing.

One limitation of this portion of the study is that we do not have a direct comparison group for these survey results. However, UCSD has a strong industry culture and relatively few undergraduate CS majors do research. In UCSD's 2018 graduating student exit survey, only 39% of CS majors reported doing undergrad research, while 85% reported having an industry internship. 24% reported plans to attend graduate school in a related field, but we have no data on how many actually applied to or enrolled in graduate school. Finally, the demographics of the students who participate in ERSF are not at all representative of the demographics of the CS department at UCSD making it a skewed comparison group.

7.3 RQ3 Qualitative Data

Open-ended question (9) also provided insight into the influence that ERSF had on students' future choices to pursue research and what, specifically, they valued about the program.

Many students wrote about how ERSF demystified research for them, sometimes opening up a new interest. For example,

Attending ERSF I experienced what doing a research is like for the first time and it sparked my interest.

–Cohort 3 participant

Sometimes this new understanding of research helped students make an informed decision that research was not the career they desired, as the following quote illustrates:

Participating in ERSF helped me to gain experience and knowledge that I otherwise would not have been able to at that early stage of my career. I had originally never considered and was completely uninterested in doing grad school for CS, but after the program I seriously considered it for a while before I ultimately chose to take a different career path.

–Cohort 1 participant

Whether participants ended up pursuing a career in industry or continuing to graduate school, many valued the specific skills and knowledge ERSF provided. Students wrote about being exposed to new areas of computing they didn't know existed:

ERSF did introduce me to some of the areas of Computer Science I was never exposed to.

–Cohort 3 participant

My ERSF experience helped me gain more exposure with machine learning... which is something I would like to study further in graduate studies.

–Cohort 3 participant

Several participants valued the soft-skills ERSF helped them develop, as in the following quote:

ERSF not only helped me get early exposure to networking knowledge and big data technologies, but also made me realize how important it is to communicate well with teammates and professors/supervisors.

–Cohort 2 participant

For some students, ERSF provided a sense of belonging, connection and confidence that they did not previously have. For example, these two participants wrote about how this confidence and sense of belonging helped them pursue further research opportunities.

It allowed me the opportunity to get an introduction to research in the CS community and made me feel as though I belonged in this field. I learned a lot and ERSF was a huge reason why I applied to and got accepted to graduate schools. –Cohort 3 participant

ERSF made me more open to research, and I was less scared to get involved with it with another professor and topic. –Cohort 1 participant

Other students specifically cited the team-based and community-focused nature of the program as valuable, for example:

The most valuable part [of ERSF] was working on CS research with my group and learning from them.

–Cohort 3 participant

While many structured undergraduate research programs provide some kind of cohort and community-building experience, this response and others like it highlight importance of the group-based structure and close mentoring to build a shared experience.

Finally, although retention rates of ERSP participants and Finalists were not significantly different, for at least one participant, ERSP was a significant reason for staying in CS:

ERSP helped me by motivating me to finish my major. Before I started ERSP, I was not sure if computer science was the right major for me and even briefly considered dropping out of the major. ERSP motivated me to finish my degree. —*Cohort 2 participant*

8 DISCUSSION

The data and analysis presented in this study give us a broad perspective on how ERSP may affect its participants. In this section, we contextualize these results and discuss the limitations of our work.

8.1 Increased opportunities for research

ERSP has drastically increased the opportunities at UCSD for second-year undergraduates, particularly those from groups underrepresented in computing, to engage in research projects. The scale of this program would not be possible without the group-based, dual mentoring structure of ERSP.

It was not obvious that this group-based, split-mentoring approach to scaling early undergraduate research would work. Fortunately, this early group-based exposure to research appears to have many benefits for participants. Students who participated in ERSP had higher grades and retention than students in a general control group, and our analysis suggests that they might have higher grades than students who were deemed eligible for ERSP but did not participate. Compared to students who complete more traditional REUs, recent ERSP participants appear to feel more confident in their ability to succeed academically and have a higher sense of belonging. From open-ended survey responses, ERSP increased students' understanding of research and the field of CS, and for many of them it encouraged them to pursue further research opportunities and apply to graduate school.

8.2 The evolution of ERSP

Over its five years, ERSP has evolved and improved. As one example, mentors are trained more carefully, and starting with cohort 3, projects are defined before the program begins and refined until they are appropriate for early undergraduate students. They are also published, and students are allowed to rank their interest in the specific projects rather than just in general areas of CS.

These improvements are reflected in some of the data presented here. For example, the analysis in Section 6 showed no significant benefits for ERSP participants until cohort 3, which could be due to the improvements in the program structure. Also, the only negative comment about ERSP we received in the open-ended feedback to the survey in Section 7 came from a cohort 1 participant.

ERSP will continue to evolve as it spreads beyond its initial context at UCSD. UC Santa Barbara, University of Illinois–Chicago and Stanford are all developing and launching a version of ERSP. ERSP's growth to these different contexts, will open more opportunities to study the impact of its structure on a wider scale.

8.3 Beware selection bias

The use of two control groups for our grade and retention analysis allowed us to examine the substantial effect that selection bias can play in an analysis such as the one presented here. We matched our comparison group on several student characteristics, but even so we saw a big difference between the Comparison group and the Finalists. In both grades and retention, the difference between ERSP participants and Finalists was fairly small and in the case of retention not significant. A possible explanation for this trend is that a student who is sufficiently motivated to submit a high-quality application for ERSP will seek out (and get admitted to) other opportunities that produce similar gains. However, even if this is the case, ERSP still has value in providing more such opportunities for these highly motivated students.

The Finalist control group presented here provides a gold-standard matched group: the Finalists are identical to ERSP admitted students but randomly selected out. Program designers should consider arranging their selection process to produce such a comparison group to help evaluators measure the true impact of the program.

8.4 Threats to validity

Like any quantitative study, we face threats to validity pertaining to the size and composition of the treatment group and the comparison group(s). In the quantitative study in Section 5, the difference between the Finalists and the ERSP participants is barely significant at the 0.05 level. This result could be because the Finalist group is small, but we must interpret the difference with caution and not conclude too much. Another threat comes from the way we handled admitted students who did not complete ERSP. Students who dropped out of the program had a lower GPA than those who completed it (though not quite significantly so), so excluding those students might have introduced bias in the ERSP group.

For the surveys in sections 6 and 7, response bias may have influenced the results. While our response rates were high, it is possible that those who responded to the survey were more positive about the program, research and graduate school, causing the percentages and values reported to be higher than their true values.

Finally we note that most of the work presented here is correlational. While we attempted to build comparison groups as similar to the ERSP group as possible, this is not a controlled experiment so we cannot conclusively say that participation in ERSP caused any of the effects presented here.

9 CONCLUSION

The work presented here helps validate the ERSP structure and adds to our collective understanding of the value of early undergraduate research in CS. While there is much more to be studied within ERSP specifically and about early undergraduate CS research generally, we hope that these results will encourage others to develop similar programs to engage a larger and more diverse group of early undergraduates in CS research.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. DUE-1821521.

REFERENCES

- [1] A. Bandura. 1977. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84, 2 (1977), 191–215.
- [2] L. Barker. 2009. Student and Faculty Perceptions of Undergraduate Research Experiences in Computing. *Trans. Comput. Educ.* 9, 1, Article 5 (March 2009), 28 pages. <https://doi.org/10.1145/1513593.1513598>
- [3] Karen W. Bauer and Joan S. Bennett. 2003. Alumni Perceptions Used to Assess Undergraduate Research Experience. *The Journal of Higher Education* 74, 2 (2003), 210–230.
- [4] Bethany Bowling, Heather Bullen, Maureen Doyle, and John Filaseta. 2013. Retention of STEM majors using early undergraduate research experiences. In *Proceeding of the 44th ACM technical symposium on Computer science education*. ACM, 171–176.
- [5] Barbara Ericson. Accessed August 20, 2017. AP Data for the United States 1998–2016. <http://home.cc.gatech.edu/ice-gt/595>. (Accessed August 20, 2017).
- [6] S. R. Gregerman, J. S. Lerner, W. von Hippel, J. Jonides, and B. A. Nagda. 1998. Undergraduate Student-Faculty Research Partnerships Affect Student Retention. *The Review of Higher Education* 22, 1 (1998), 55–72.
- [7] Carlos G. Gutierrez, Linda M. Tunstad, Anthony Fratiello, and Scott L. Nickolaissen. 2003. Undergraduate research participation increases minority retention and success in chemistry. In *Abstracts of Papers, 225th ACS Nat'l Mtg*.
- [8] R. W. Lent, S. D. Brown, and K. C. Larkin. 1984. Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology* 31, 3 (1984), 356–362.
- [9] R. W. Lent, S. D. Brown, and K. C. Larkin. 1986. Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology* 33, 3 (1986), 265–269.
- [10] Marcia C. Linn, Erin Palmer, Anne Baranger, Elizabeth Gerard, and Elisa Stone. 2015. Undergraduate research experiences: Impacts and opportunities. *Science* 347, 6222 (2015). <https://doi.org/10.1126/science.1261757> arXiv:<http://science.sciencemag.org/content/347/6222/1261757.full.pdf>
- [11] David Lopatto. 2004. Survey of Undergraduate Research Experiences (SURE): First Findings. *Cell Biology Education* 3, 4 (2004), 270a–277.
- [12] David Lopatto. 2007. Undergraduate Research Experiences Support Science Career Decisions and Active Learning. *Cell Biology Education* 6, 4 (2007), 297–306.
- [13] Jane Margolis. 2008. *Stuck in the Shallow End: Education, Race, and Computing*. The MIT Press.
- [14] Jane Margolis and Allan Fisher. 2003. *Unlocking the Clubhouse: Women in Computing*. The MIT Press.
- [15] Engineering National Academies of Sciences and Medicine. 2017. *Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/24622>
- [16] Ama Nyame-Mensah. 2015. Strengthening the computing research pipeline through minority participation: The case for undergraduate research experiences. In *2015 Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*. IEEE, 1–7.
- [17] J. Pearl. 2009. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, Chapter Understanding Propensity Scores.
- [18] Matea Pender, Dave Marcotte, Mariano Sto. Domingo, and Kenneth Maton. 2010. The STEM Pipeline: The Role of Summer Research Experience in Minority Students' Graduate Aspirations. *education policy analysis archives* 18 (2010).
- [19] P. R. Rosenbaum and D. B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70 (1983), 41–55.
- [20] Susan H. Russell, Mary P. Hancock, and James McCullough. 2007. THE PIPELINE: Benefits of Undergraduate Research Experiences. *Science* 316, 5824 (2007), 548–549.
- [21] E. Seymour, A.-B. Hunter, S. L. Laursen, and T. Deantoni. 2004. Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education* 88 (2004), 493–534.
- [22] Jill Singer and Bridget Zimmerman. 2012. Evaluating a summer undergraduate research program: Measuring student outcomes and program impact. *CUR Q* 32, 3 (2012), 40–7.
- [23] Jane G. Stout, N. Burçin Tamer, and Christine J. Alvarado. 2018. Formal Research Experiences for First Year Students: A Key to Greater Diversity in Computing?. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*. ACM, New York, NY, USA, 693–698. <https://doi.org/10.1145/3159450.3159472>
- [24] Michael F. Summers and Freeman A. Hrabowski. 2006. Preparing Minority Scientists and Engineers. *Science* 311, 5769 (2006), 1870–1871. <https://doi.org/10.1126/science.1125257> arXiv:<http://science.sciencemag.org/content/311/5769/1870.full.pdf>
- [25] Burçin Tamer and Jane G. Stout. 2016. Understanding How Research Experiences for Undergraduate Students May Foster Diversity in the Professorate. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16)*. ACM, New York, NY, USA, 114–119. <https://doi.org/10.1145/2839509.2844573>
- [26] N. B. Tamer. 2015. *REUs & Graduate School Attendance Rates*. Technical Report. Computing Research Association.
- [27] Vincent Tinto. 1993. *Leaving College: Rethinking the Causes and Cures of Student Attrition*. University of Chicago Press.
- [28] G. M. Walton and G. L. Cohen. 2007. A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology* 92, 1 (2007), 82–96.