University of Minnesota-Twin Cities Computational Methods for Discovery Driven by Big Data REU Site: Cohort 1 Evaluation Report

HEATHER M. WRIGHT
JANE G. STOUT

CERP
Computing Research Association
Evaluation

Recommended citation:
Wright, H. M. and Stout, J. G. (2016). University of Minnesota-Twin Cities Computational Methods for Discovery Driven by Big Data REU Site: Cohort 1 Evaluation Report. Available from: http://cra.org/cerp/our-services/

To learn more about CRA's Center for Evaluating the Research Pipeline, please visit our website: http://cra.org/cerp. Questions regarding this report and requests for other related documents should be sent to cerp@cra.org.

Center for Evaluating the Research Pipeline
Computing Research Association
1828 L St. NW, Suite 800
Washington, DC 20036
t. 202-266-2937
f. 202-667-1066
e. cerp@cra.org

Table of Contents

acknowledgments

CERP was created by the Committee on the Status of Women in Computing Research (CRA-W)/Coalition to Diversity Computing (CDC) Alliance through a National Science Foundation grant to the Computing Research Association (CNS-1246649). The current research was supported by NSF grant IIS-1460620. Any opinions, findings, conclusions, and recommendations are the authors' and do not necessarily reflect the views of the National Science Foundation.

CRA-W
Computing Research Association
Women

CRA
Computing Research Association

the coalition
to diversify computing

About CERP

The Computing Research Association's (CRA) Center for Evaluating the Research Pipeline (CERP) evaluates the effectiveness of intervention programs designed to increase retention of individuals from underrepresented groups in computing, namely men from underrepresented racial/ethnic groups, and women of all racial/ethnic backgrounds. More generally, CERP strives to inform the computing community about patterns of entry, subjective experiences, persistence, and success among individuals involved in academic programs and careers related to computing. For more information about CERP, visit http://cra.org/cerp/.

Executive Summary

Compared to a group of students who had not participated in an REU program, UMN REU students reported greater:

- interest in becoming a professor in a computing field
- interest in becoming a computing researcher in an industry or government lab
- aspirations to attain a doctoral degree
- involvement in extracurricular computing related activities

Introduction

The University of Minnesota- Twin Cities' (UMN) Department of Computer Science and Engineering created the "Computational Methods for Discovery Driven by Big Data" Research Experience for Undergraduates (REU) program to engage participants in big data interdisciplinary research and prepare them for a future career in the sciences. This 10-week summer research program promotes academic persistence for underrepresented groups in computing and supports activities that will motivate students for lifelong contributions to the scientific community.

The UMN REU site centers around four main objectives: (1) intellectually engage and excite participants to motivate their commitment to and pursuit of a career in the sciences, (2) increase participation in and contribution to the sciences by women and underrepresented minorities in computer science, (3) train students for sustained contribution to the sciences, particularly in computational methods for big data trans-disciplinary research, and (4) professionally prepare and mentor participants for a career in the sciences, meaning to teach participants to be effective communicators, be career savvy, and versed in the ethics of science.

Drawing students from institutions across the U.S., as well as students from UMN, the UMN REU 10-week summer research program creates a small cohort of students from geographically diverse areas. This combination of resident and non-resident participants increases diversity and develops institutional partnerships. Students, particularly from underrepresented groups, are encouraged to apply, and top-ranked students are admitted to the program. The UMN REU aims to educate students on how to formulate questions, conduct research, and communicate findings. Students participate in activities that provide technical training, career mentoring, and professional development to engage and prepare participants for future careers in the sciences and contribute to the scientific community beyond the classroom.

In this report, we focus on the experiences and aspirations of students who had participated in the UMN REU program during the summer of 2015. We compare those experiences to a sample of students who had not participated in a formal research experience during the summer of 2015. As will be seen in the Method section, the current report strives to understand how UMN REU students differ or relate to a group of students who are like them in many ways with the exception of an REU experience.

Evaluation Method

Procedure

Each fall, CERP surveys undergraduate students across the U.S. who are majoring or minoring in computing, or enrolled in computing courses. The survey assesses students' past research and professional experiences (e.g., participation in an REU program), subjective experiences in computing (e.g., sense of belonging), and aspirations for the future (e.g., interest in a professorial career in computing), and is used to evaluate REU programs such as the UMN REU site. During the fall of 2015, the first cohort of the UMN REU participants was recruited to complete CERP's annual survey of undergraduate experiences in computing. Ten UMN REU students, and 3,142 students who had not participated in the UMN REU completed the survey. Survey items and scales used for this evaluation can be found in Appendix A.

Survey Respondents

To assess the impact of the UMN REU program on students' experiences in the computing community and their career aspirations, we utilized a comparative evaluation framework. For this, we extracted students who had not completed an REU during the preceding year from our larger sample of students who had completed the CERP survey (n = 2,874). Among this group of students, we created a subsample to be compared against UMN REU students using propensity score matching. Propensity score matching is an analytic technique that "matches" individuals in a treatment group (e.g. UMN REU students) to individuals from a comparison group (e.g., non-REU students) who are as comparable as possible on a set of relevant individual-level characteristics [1,2]. Thus, propensity score matching controls for the role of student-level variables (e.g., year in college) that might otherwise explain outcome variables (e.g., self-efficacy) rather than the "treatment". Non-REU students were matched to UMN REU students on five variables: academic class, age, Carnegie institutional classification of home institution's research rigor, major, and underrepresented status (Asian and/or White men versus men from minority racial/ethnic groups and women). Appendix B shows students' demographics for the two matched samples.

A Note About the Analyses

The sample sizes for the UMN REU group and the non-UMN REU group were small (n = 10 for each group), rendering the reliability our findings potentially unstable. As such, we urge readers to interpret the following analyses with caution. Future evaluation work conducted by CERP for the UMN REU program will examine data across multiple cohorts of participants, yielding a larger sample size, and more reliable analyses.

Evaluation Results

Dependent measures were either aggregated across multiple items when reliability was strong, or measured using a single survey item. See Appendix A for the reliability statistics for aggregate measures. We compared UMN REU vs. non-REU mean scores for Likert style measures using one-way Analyses of Variance (ANOVAs). We used Pearson Chi Square tests to compare frequency data (i.e., categorical responses) across the two student groups. We used a two-tailed test with a cut-off alpha criterion of $p \leq .05$ to determine statistical significance for each analysis.

Aspirations and Involvement

Aspirations to pursue computing-related science careers, and highest degree intentions. Survey respondents were asked about their academic and career interests and aspirations. As seen in Table 1 on page 10, UMN REU students reported stronger interest in becoming a computing researcher in the academy, $F(1,18) = 5.64$, $p \leq .05$, and in industry/government, $F(1,18) = 7.23$, $p \leq .05$. Note that all students' interest in the professorate was relatively low (i.e., below the scale midpoint), as was interest in becoming a high school computing teacher. No other group differences in career aspirations emerged.

We also found UMN REU students were more interested in pursuing a Ph.D. than non-REU students, $p \leq .05$. By contrast, Non-REU students' highest degree intentions were more likely to be a B.S. than UMN REU students, $p \leq .05$.

Involvement in Computing-Related Activities. We conceptualize involvement in computing-related extracurricular activities as an index of engagement in computing. We measured students' involvement with 10 different computing activities during the year preceding the survey. Percentages of students who had participated in each activity within each group are listed in Table 2 on page 11. We also computed a mean score for the number of activities students were involved with ranging from 0 – 10. As seen in the bottom row of Table 2, UMN REU students were more involved with the 10 listed computing related activities than Non-REU students, $F(1,18) = 5.05$, $p \leq .05$.

Table 1. Aspirations to pursue computing-related science careers, and highest degree intentions.

	UMN REU	Non-REU
Aspirations to pursue computing-related science careers		
Scale anchors ranged from (1) *Very uninterested* to (5) *Very interested*		
College/University professor in computing field	2.90	1.70*
Computing researcher in industry or government lab	4.00	2.60*
High school computing teacher	1.50	1.30
A non-research position in the computing industry	4.10	3.60
Interdisciplinary research position (applying computing knowledge)	3.70	3.00
Interdisciplinary non-research position (applying computing knowledge)	3.30	3.60
Computing related Entrepreneur	3.50	3.60
Non-computing career	2.60	2.45
Highest degree intentions		
B.S.	20%	70%*
M.S.	20%	10%
Ph.D.	50%	0%*
Plan for highest degree to be in a computing field	90%	80%

*$p \leq .05$; comparison against UMN REU participants.

Note. Values for "Career-path interests" indicate mean score within each group. Values for "Highest degree intentions" indicate percentages within each group.

Table 2. Involvement in Computing-Related Activities

	UMN REU	Non-REU
Visiting lectures in your department related to computing	60%	30%
Computing-related student groups	60%	20%
Computing-related contests (hacking, robotics competitions, etc.)	30%	30%
Computing-related online social networking (listservs, etc.)	70%	20%
Professional societies related to computing	30%	10%
Technical conferences related to computing	40%	10%
Outreach to K-12 students related to computing	30%	10%
Training or workshops in computing (other than conferences)	20%	0%
Summer institutes or short courses (other than summer research programs)	20%	10%
Study support in computing (e.g., Supplemental Instruction (SI), pair programming)	30%	0%
Mean number of activities listed above	3.90	1.40*

*$p \leq .05$; comparison against UMN REU participants.

Note. Respondents could select more than one option. Values represent percentage within each group. "Mean number of activities" ranges from 0 – 10, where 0 indicates involvement with none of the listed activities, and 10 indicates involvement with all 10 listed activities.

Sense of Identity and Support

As seen in Table 3 below, students' level of belonging, computing identity, mentor support, and self-efficacy did not differ across student groups. Note students' level of mentor support was below the midpoint of the scale, suggesting room for improvement in students' perceived support from mentors.

Table 3. Computing identity, belonging, support, and self-efficacy

	UMN REU	Non-REU
Computing Identity	4.08	4.00
Belonging	3.87	3.50
Mentor Support	2.77	2.03
Self-efficacy	4.17	3.77

Note. Values indicate mean score within each group.

Beliefs about Computing Careers

We measured the degree to which students believed they would be able to do the following with a computing career: serve humanity, be in a position of influence in society, and spend time with family. Compared to non-REU students, UMN REU students' beliefs about computing careers were not statistically different, as seen in Table 4 below.

Table 4. Beliefs about Computing Careers

Scale anchors ranged from (1) *Low* to (5) *High*	UMN REU	Non-REU
Opportunity to serve humanity	4.30	3.60
Able to be in a position of influence in society	3.90	3.60
Able to spend time with family	3.60	3.20

Note. Values indicate mean score within each group.

Summary and Conclusion

We found that UMN REU students reported stronger interest in becoming computing researchers in academia and in industry/government. In addition, we found UMN REU students were more interested in pursing Ph.D. programs compared to non-REU students. These findings are consistent with the program goals to foster academic persistence in the computing research pipeline, and to encourage and motivate students towards careers in the sciences.

We think it is noteworthy that UMN REU students reported higher levels of self-efficacy compared to Non-REU students, though this effect was not significant (see Table 3). This finding corresponds with the REU's goal to prepare students professionally and mentor students for a career in science. Also noteworthy, though not statistically significant, was UMN REU students' perceived higher ability to serve humanity in a computing-related science career than Non-REU students (see Table 4). This finding coincides with the REU's focus on the interdisciplinary nature of big data, and by extension, the social applications of computing. Given that these findings are consistent with the goals of the UMN REU program, we expect that, with greater sample sizes, significant group differences would have emerged. Indeed, a post hoc power analysis indicates that we would have needed a sample of size of 30-35 students per comparison group to find that these effects were significant. Future evaluation reports from CERP will merge data from multiple cohorts in order to boost statistical power, and provide a more rigorous evaluation of the UMN REU program.

References

Bai, H. (2015). Current issues on propensity score matching: Matching with/without replacement, common support, and sample ratio. In W. Pan, & H. Bai (Eds.), *Propensity score analysis: Fundamentals, developments and extensions.* New York, NY: Guilford Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Appendix A: Survey Items

Our measures included single survey items as well as constructs that combine multiple survey items into a single measure. We determined reliability for multi-item constructs using Cronbach's alpha (α). Alpha levels ≥ .70 are considered acceptable. Below are all the survey items used this report, as well as alpha levels for constructs.

Aspirations and Involvement

Aspirations to pursue computing-related science careers

How interested are you in having the types of jobs listed below after you finish your highest degree?

- College/University professor in computing field
- Computing researcher in industry or government lab
- High school computing teacher
- A non-research position in the computing industry
- Position applying computing research to another area (e.g. digital media, support of research in medicine or other sciences)
- Non-research position applying your computing knowledge in another area (e.g. business applications, government)
- Entrepreneur (computing related)
- Non-computing career

Very disinterested (1)	Somewhat disinterested (2)	Neither disinterested nor interested (3)	Somewhat interested (4)	Very interested (5)

Highest degree intentions

What is the highest degree you plan to attain?

- Bachelor's degree
- Master's degree
- Doctoral degree

Plan for highest degree to be in a computing field

Survey respondents were coded as planning for highest degree to be in a computing field if any of the following were selected:

In which field do you plan to attain that degree? Please select all that apply.

- Computer Science
- Computer Engineering or Electrical and Computer Engineering
- Computing Information Systems or Information Systems
- Other computing major; please specify: _____

Survey respondents were coded as planning for highest degree to be in a non-computing field if any of the following were selected:

In which field do you plan to attain that degree? Please select all that apply.

- Math/Applied Math
- Business or Law
- Life/Health Sciences
- Interdisciplinary, please specify areas: _____
- Other science or engineering non-computing major; please specify: _____
- Uncertain
- Other non-computing major; please specify: _____

Involvement in computing-related activities

During the past year, were you involved in any of the following groups or activities?

- Visiting lectures in your department related to computing
- Computing-related student groups
- Computing-related contests (hacking, robotics competitions, etc.)
- Computing-related online social networking (listservs, Facebook groups, etc.)
- Professional societies related to computing
- Conferences related to computing
- Outreach to K-12 students
- Summer institutes or short courses (other than summer research programs)
- Study support in computing (e.g. Supplemental Instruction (SI), pair programming)
- Trainings or workshops in computing (other than conferences)

No, I have not participated in this group activity (0) Yes, I have participated in this group activity (1)

Sense of Identity and Support

Belonging, $\alpha = .90$

How much do you agree or disagree with the following statements?

- I feel like I 'belong' in the computing community.
- I feel like an outsider in the computing community *(reverse-coded)*.
- I feel welcomed in the computing community.

Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)

Mentor support, α = .93

To what extent do you have a mentor who...

- helps you improve your computing skills.
- shows compassion for concerns and feelings you discussed with them.
- shares personal experiences as an alternative perspective to your problems.
- explores career options with you.

Not at all (1)	A little (2)	Somewhat (3)	Quite a bit (4)	Very much (5)

Self-efficacy, α = .84

I am confident that I can...

- find employment in my area of computing interest.
- become a leader in the field of computing.
- win a computing-related contest (e.g., programming contest, robotics contest, hackathon).
- get admitted to a graduate computing program.
- complete my undergraduate degree in computing.
- quickly learn a new programming language on your own.
- clearly communicate technical problems and solutions to a range of audiences.

Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)

Beliefs About Computing Careers

In your opinion, to what extent would a career in computing allow you to do the following?

- Serve humanity
- Be in a position of influence in society
- Spend time with family

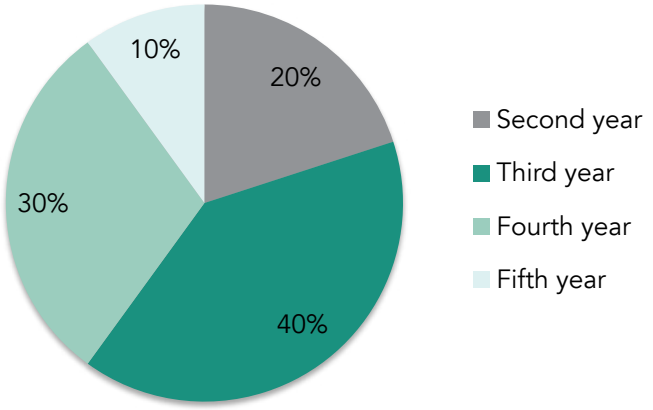
Not at all (1)	A little (2)	Somewhat (3)	Quite a bit (4)	Very much (5)

Appendix B: Sample Characteristics

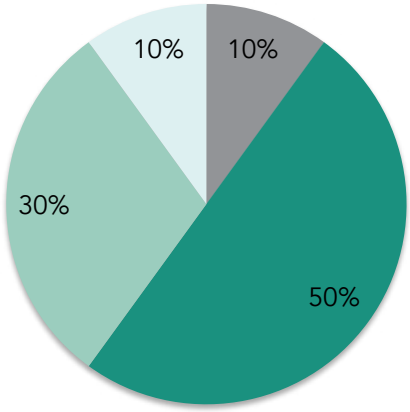
This section displays demographics of UMN REU participants (n = 10) and matched sample of students who had no REU experience (n = 10).

Academic class

UMN REU Students

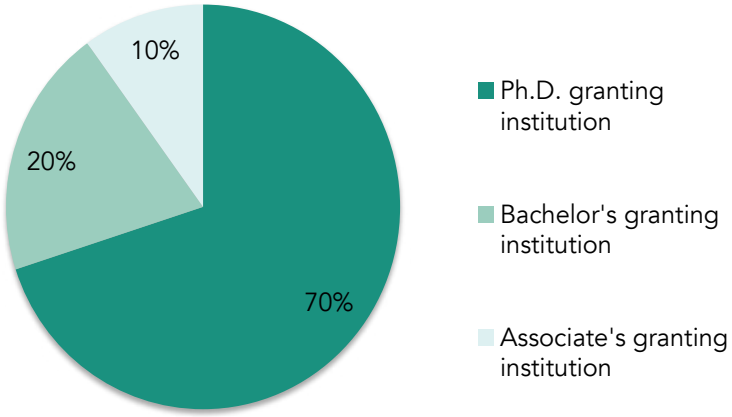


Non-REU Students

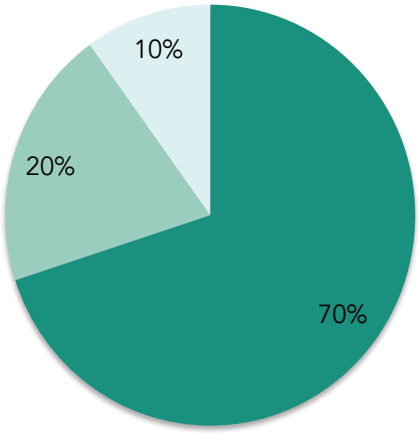


Institution type

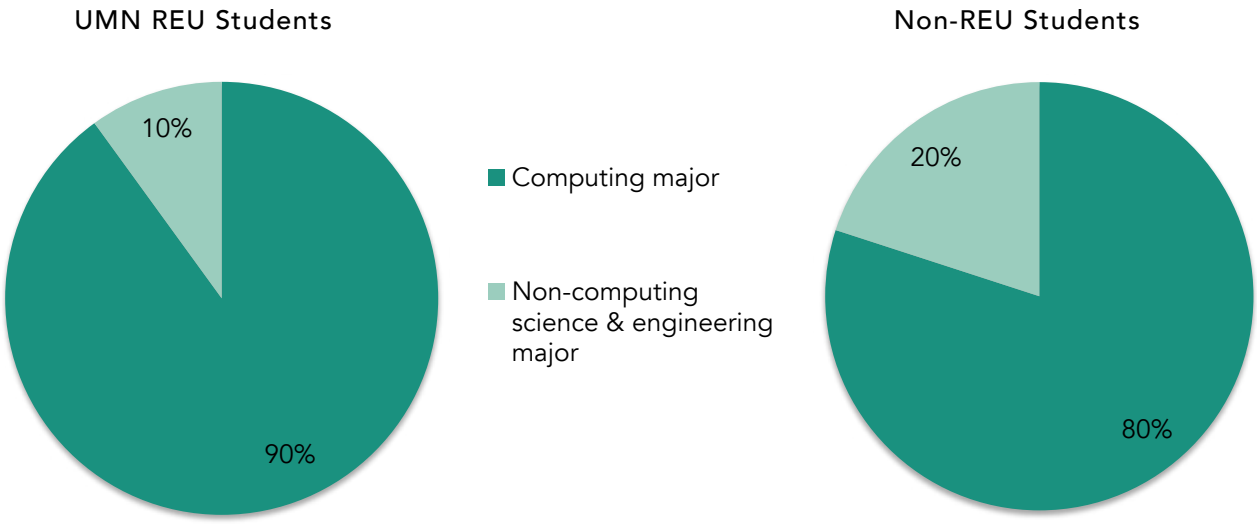
UMN REU Students



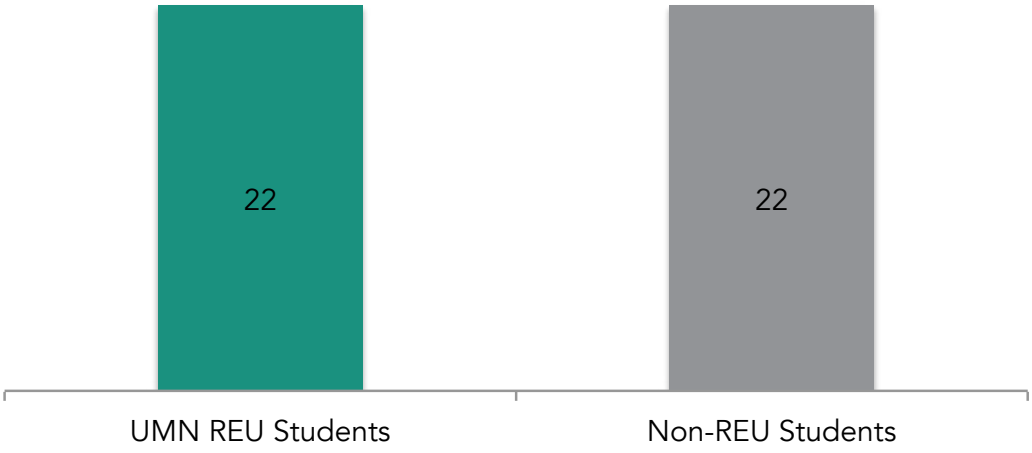
Non-REU Students



Major



Median age



Underrepresented status