Holistic Summer Undergraduate Research Program in High Performance Computing Research

Tyler Miller
Dept. of Psychology
South Dakota State University
Brookings, SD, United States
Tyler.Miller@sdstate.edu

Jung-Han Kimn
Dept. of Mathematics & Statistics
South Dakota State University
Brookings, SD, United States
Jung-Han.Kimn@sdstate.edu

Stephen P. Gent
Dept. of Mechanical Engineering
South Dakota State University
Brookings, SD, United States
Stephen.Gent@sdstate.edu

ABSTRACT

This study investigated the effectiveness of a National Science Foundation-sponsored Research Experience for Undergraduates Site in High Performance Computing (HPC) at South Dakota State University. In the 2018 program, ten first- and second-year students majoring in science, technology, engineering and math (STEM) disciplines were engaged in an intensive ten-week program on the use and application of HPC in theoretical and applied research. The participants were engaged in an interdisciplinary program and were mentored by faculty and graduate students as they completed their project. Participants were provided instruction on the use of HPC, preparing posters and presentations, the Responsible Conduct of Research, and a variety of interpersonal skills. Feedback on the program was evaluated with daily journal entries prepared by the participants, Computer & Information Science & Engineering (CISE) pre- and post-survey responses, one-on-one interviews, and the Torrance Tests of Creative Thinking. Results of the 2018 cohort indicate that participants' attitudes toward their project discipline (e.g., engineering or biological sciences) are favorable. Participants' self-reported Research Skills also increased from the beginning to the end of the program. However, results suggest that their favorable attitudes decreased from the beginning to the end of the program. One hypothesis of why this occurred is that participants had unrealistic ideas about the time, effort, or duties of high-level research in math, engineering, and/or biological science disciplines. After the ten-week program, participants' views were more closely aligned with the reality of research work in the area.

CCS CONCEPTS

• CS→Computer Science; • Cybertraining→training on using cyberinfrastructure; • HPC→high performance computing

KEYWORDS

High Performance Computing, Undergraduate Research Experiences, Project-Based Learning, Mentorship, Responsible Conduct of Research

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 ACM 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. PEARC '19, July 28-August 1, 2019, Chicago, IL, USA © 2019 Association for Computing Machinery. ACM ISBN: 978-1-4503-7227-5/19/07...\$15.00
<https://doi.org/10.1145/3332186.3332215>

ACM Reference format:

T. Miller, J.-H. Kimn, and S. Gent. 2019. A Holistic Summer Undergraduate Research Program in High Performance Computing. In *Proceedings of ACM Practice & Experience in Advanced Research Computing Conference Series, Chicago, Illinois USA, July 2019 (PEARC'19)*.

1 INTRODUCTION

Having a workforce that is well prepared in High Performance Computing (HPC) is essential for the 21st century. Despite the ubiquity of computers in our society, much of the training required to become proficient in computationally-intensive professions (modeling & simulation, analytics, etc.) requires a larger and more diverse HPC workforce in academic, industry, government and other organizations [1, 2].

Instruction in parallel and distributed computing has traditionally been relegated to a few isolated courses, taught primarily in the context of scientific computing, distributed systems, or computer networks [3]. The effective application of high performance computing (HPC) is a skill increasingly demanded in a widening variety of fields including engineering, mathematics, statistics, physics, biology, plant science, biochemistry, and many others. The program described below was designed to meet this demand. The objectives of the program were to: 1) increase student participation in conducting research from under-represented populations, small rural colleges, tribal colleges, community colleges, and veterans returning from active duty, 2) enhance participants' skills and

capabilities to conduct innovative and meaningful research using HPC, 3) strengthen participants' capabilities for learning and thinking independently within an interdisciplinary collaborative environment, 4) convey the importance of conducting research with the utmost integrity, 5) improve student preparation for their professional careers through integrated professional development activities, and 6) encourage participants to pursue graduate degrees in STEM programs.

This project encompassed several STEM disciplines (mathematics and statistics, mechanical engineering, civil engineering, chemistry and biochemistry, and plant science), which use similar simulation tools, techniques, and requirements but have significantly different applications. Student participants engage in research as they learn how high-fidelity models represent physical phenomena; how these models are implemented within HPC; and how results are analyzed, interpreted, and utilized to gain insight into theoretical and applied problems. Each student worked closely with a graduate assistant mentor on a daily basis.

2 EXPERIMENTAL AND COMPUTATIONAL DETAILS

2.1 Recruitment and Selection

This REU site recruited ten undergraduate students out of an applicant pool of 104 students after their first or second year of college for our ten-week summer program. We recruited a diverse pool of talented and traditionally underrepresented students as well as veterans returning from active duty. More than 40 percent of participants were recruited from smaller liberal arts colleges and two-year technical colleges that have limited or no formal research programs.

The program encompassed a shared community of computational researchers at SDSU organized around the use of HPC in their respective STEM disciplines (mathematics and statistics, engineering and applied science, as well as fundamental science disciplines, including biology and chemistry). Each HPC project was overseen by a faculty mentor in their established laboratory. Also, each student worked closely with a graduate assistant mentor on a daily basis.

2.2 Nature of the Research Experience

The core of this project is research activities in three focus areas encompassing HPC/Big Data and holistic training in research integrity and professional development: 1) Engineering and Technology, 2) Computational Mathematics & Statistics, and 3) Biological Sciences. After the selection process, each student participant was in contact with their faculty mentor via email and telephone prior to the summer program where the faculty mentors provided a list of recommended reading materials and tailored projects for each student. During the first three days of the program, participants received training in theoretical foundations, numerical methods, and software appropriate to their projects. Examples of the topics covered in the first portion of the HPC Education were as follows.

Overview of HPC

- What is High Performance Computing?
- HPC and Big Data applications to research activities by research faculty mentors

- General advantages of running computing jobs on a remote server or in the cloud

Discussion and hands-on activities: What can a cluster do that your laptop cannot?

- Example of scalable, distributed, parallel computing: simulation of a reactive flame

- How to automatically run the same program many times with different inputs

The specific HPC projects were selected based on estimated completion time, level of difficulty, and individual student's interest. Crosscutting projects were arranged in the prior spring for student scholars to experience interdisciplinary collaboration. Under the mentorship of their respective faculty members, student participants prepared technical presentations and conference-quality publications based on their research. A series of discussions and case studies on research integrity and HPC is an integral component of this program. Table 1 below outlines the activities done by the cohort as a group.

Table 1: Weekly Timeline of Student Activities

Week	Student Activities
1	Orientation, Introduction to Topics, and Literature Review
2	Slide Presentations, Initial HPC Education, and Research Integrity I
3	Initial Simulations and Further HPC Education
4	Set up simulations and Research Integrity II
5	Check Simulation runs and Midterm Presentation
6	Draft of Final Manuscript and Research Integrity III
7	Model Diagnostics and Simulations Review
8	Wrap-up technical portions of projects
9	Final Manuscript and Research Integrity IV
10	Final Presentations

Project Synopses: Brief summaries of the projects completed by the 2018 cohort are listed below to provide context on the nature of the projects.

Project 1: "Development of CFD models to predict performance of additively manufactured plastic heat exchangers." 3D printed heat exchangers allow for creative designs that are not possible with traditional manufacturing methods. Fluid flow and heat transfer in 3D printed heat exchangers was simulated using computational fluid dynamics (CFD) modeling. Three experimentally tested heat exchangers were used to validate the simulations. Next, a new design using trapezoidal shaped tubes in a wavy pattern was developed. In the future, CFD modeling could complement the use of experimental trials, reducing costs and resources used.

Project 2: "Towards High Fidelity Simulations of Biofilm Fluid Dynamics." As part of combining the Navier-Stokes and Cahn-Hilliard equations, we investigated numerical implementations of the Flory-Huggins free energy density equation and its first variation. The completed project can be used to model biofluids and more.

Project 3: "Analysis of Spatial Agricultural Data: Explaining Variation in Corn Yield Using Multiple and Spatial Regression." The objective of this project was to develop a statistical

methodology to find a global model and local models using multiple and spatial regression. Model set up involved combining two diverse spatial data sets and writing an R function to find a multiple regression model to explain yield variation. Key finding was the successful combination of both data sets. This methodology can be applied to other data sets and used to model other response variables.

Project 4: "Modeling Endovascular Treatment of Unruptured Intracranial Aneurysms using Computational Fluid Dynamics." A variety of endovascular treatments for intracranial aneurysms are available, but there is limited data comparing the effectiveness of such treatments from a fluid dynamics perspective. Computational fluid dynamics was used to compare the effectiveness of standard coiling, hydrogel coiling, and flow disruption treatments. Results from simulations of blood flow through the models were compared to determine which treatments were most effective for each aneurysm test case.

Project 5: "Sunflower Phomopsis Stem Canker Image Analysis." South Dakota is the nation's largest producer of sunflowers, because of this the crop needs protected. Phomopsis stem canker--a prolific disease--needs a more efficient diagnosis. The student worked on image analysis interested in the physiological composition related to the growth of sunflowers to couple the results with a molecular composition gathered from a Raman spectrophotometer. The student utilized the marker controlled watershed method and K-means clustering for image segmentation to analyze scanning electron microscope (SEM) images.

Project 6: "Grape Berry Ripening Signature Gene Expression Analysis." Frontenac and Marquette are cold climate grape hybrids. To produce good wines, it is critical to characterize their ripening profile and identify optimal biomarkers. Gene expression was analyzed at different maturity stages to identify key genes and pathways linked with flavor development. Using results from data exploration, the cultivars showed distinct signature gene expression during the ripening process. In future research, combined genetic, sensory, and chemical analyses will aid in identifying improved harvest biomarkers.

Project 7: "Unreplicated Strip Trial using Monte Carlo Simulation." The objective of this project was to develop a valid statistical methodology to determine if a treatment effect is significant for an unreplicated strip trial. Monte Carlo Simulation is used on this project by using geostatistical methods based on control strips adjacent to the trial. An open source software named R was used for this project. The main finding was the empirical type one error rate using a paired t-test was 0.4 not 0.05.

Project 8: "Towards a Biofilm Model using HPC." Biofilms are a group of microorganisms that adhere to each other and a surface. Biofilms are studied for a variety of reasons. They are studied for use in bioremediation, and according to the Centers for Disease Control (CDC) they are responsible for most chronic medical infections. The research team used the Cahn-Hilliard equation because it models how fluids can mix and separate into pure states. Utilizing the data structures contained in C and the Portable, Extensible Toolkit for Scientific Computing (PETSc) libraries we

crafted a parallel implementation of the Cahn-Hilliard equation to model the biofilm.

Project 9: "Simulating Coil Embolization Treatments of Intracranial Aneurysms using Computational Fluid Dynamics." Coil embolization is a common endovascular procedure used for treating intracranial aneurysms, which are estimated to occur in 4% in the population. In this study, a commercially available computational fluid dynamics program was used to simulate embolization techniques, standard coiling and stent-assisted coiling (SAC), in simplified vessels. On average, standard coiling and SAC reduced blood volume flow into the aneurysm by 50% and to over 60%, respectively.

Project 10: "Parallel Implementation of AC Optimal Power Flow and Time Constrained Optimal Power Flow." The objective of this project is to develop a scalable and parallel solution to the Alternating Current Optimal Power Flow (ACOPF) and Time Constrained Optimal Power Flow (TCOPF) problems, which contribute to the economic efficiency and reliability of the electric power system. This project uses C and PETSc libraries, particularly DMNetwork and Krylov Methods (KSP). The solution's effectiveness will be evaluated by comparing computation time and number of processors used.

HPC/ Big Data Education: Each of the participants were involved in at least one of three research programs: 1) Theoretical and Applied Mathematics & Statistics, 2) Engineering and Technology, and 3) Biological Sciences. Resources for HPC and research computing introduction and training/examples were provided through the SDSU University Network and Research Computing group (UNRC) and research faculty advisors. The UNRC group has two system administrators and two science domain specialists. The domain specialists were specifically tasked to work with users directly, to develop, install and test HPC applications for campus users and to perform training and outreach to current and prospective users. The domain specialists were Ph.D. level scientists with substantial previous research experience. The domain specialist directly working with this REU site was in the area of computational science/engineering. Research faculty advisors were involved in the sequence of HPC training and outreach activities, as outlined below. All were present for portions of the first presentations on Days 1 & 2 and assisted for parts of the Weeks 1 & 2 sessions, as noted. Beyond the first two weeks, faculty mentors offered additional presentations to illustrate applications more in-depth, which are more specific to individualized projects.

Responsible Conduct of Research: In addition to HPC Education, a series of four Responsible Conduct of Research (RCR) sessions were facilitated for the cohort to discuss and learn about research misconduct in the context of computing and data management. Topics for RCR included: 1) Overview of Research Integrity, Compliance, and Responsible Conduct of Research, 2) Case studies in RCR, 3) Research Misconduct and consequences in the context of simulation- and computational-based research and professional societies' codes of ethics, and 4) data fabrication, falsification, manipulation, and related topics.

Professional Development: A successful STEM researcher not only must have a strong technical background, but is also expected to be

a well-rounded professional. Higher-order thinking skills, including effective problem solving, critical thinking, and collaboration were integrated within the program through instruction in productive reflective processes and reinforced through daily written logs and reflections. A variety of personal, interpersonal, and leadership skills, including technical communication, time management, networking, and continuous improvement were integrated within the program through complementary hands-on/minds-on activities. These were facilitated through a variety of instruments, such as paper critiques, informal discussion, and case studies. As an example of enriching communication skills, all student participants prepared initial, midterm, and final presentations and a final written report at set milestones in the program. The initial and midterm presentations were presented to an audience of their peers, graduate mentors, and faculty mentors, and final presentations were open to the public. Each student participant has the opportunity to present their research at regional/national conferences. Previously, these conferences included the ASME Design of Medical Devices Conference (national), the IEEE Computing and Communication Workshop and Conference (national), and the South Dakota EPSCoR Undergraduate Research Symposium (statewide), to name a few [4, 5, 6].

2.3 External Evaluation and Assessment

The external evaluation and assessment was conducted by the first author in collaboration with the second and third authors, who also served as Co-PIs on the grant funding for the REU Site. The evaluation included three principal methods to assess the program objectives. First, we used the CISE REU Toolkit “A La Carte Student Survey” which contained questions from validated surveys and some questions that were specific to the REU [7]. Participants’ responded to questions regarding their self-efficacy, intentions toward graduate school, attitudes about the REU project discipline, help seeking (coping), grit, research skills, scientific leadership, scientific identity, faculty mentor, graduate student mentor, and the program overall.

Some of the measured constructs are self-explanatory. For example, one needs no explanation for a respondent’s intentions toward graduate school, their attitudes about the REU project discipline, and research skills. However, some constructs need explanation. For example, for our purposes the help seeking (or coping) measure described a person’s thoughts, actions, and strategies after failing a hypothetical exam. Does the person ruminate on the failure or think about steps to prevent the outcome in the future? Grit is indicative of perseverance and passion for long-term goals. Scientific leadership is a person’s knowledge about effective teamwork and confidence in their ability to lead or effectively participate in a team. Scientific identity can be summarized as a person’s endorsement of the following survey item: “I am a scientist”.

Participants completed this survey twice – once in the first week and once again in the final week of their REU experience. Questions that referred to the participants’ experiences with the program overall, their mentor, and general comments only appeared on the final week (post) survey.

Second, we used the Torrance Tests of Creative Thinking (TTCT) [8]. Creativity, broadly defined, is a novel and useful method to

solve a problem. Assessing creativity was a strategy to evaluate multiple program objectives. For example, the program objective to “Enhance undergraduates’ skills and capabilities to conduct innovative and meaningful research” directly aligns with the creative process. Like the CISE survey, participants completed the TTCT in the first week and final week of the program. The TTCT uses verbal and figural assessments to measure 5 components of creativity, 1) Fluency, 2) Flexibility, 3) Elaboration, 4) Originality, and 5) Openness. The TTCT is a standard in the literature and has been featured in many studies about creativity in elementary-aged participants and a large percentage of studies with college-aged and adult participants. The TTCT is also highly reliable, has national norms for reference, and has two versions (Forms A and B) which is ideal for pre- and post-testing. Finally, multiple sources of qualitative data were collected from REU site participants. Student participants kept a daily log of the activities and responded to open-ended questions that encouraged them to reflect on their experiences. Furthermore, throughout the 10-week period, the first author met with 9 of the 10 participants for an unstructured interview on their progress, preparation and other topics.

3 RESULTS AND DISCUSSION

3.1 Program Objectives

Regarding the program objectives outlined in Section 1, the program was successful. Half of the participants identified as female, one identified as multiethnic, one as Native America/American Indian, one as Asian, and seven as Caucasian/White. Two participants were from community or technical colleges and one participant was a first-generation college student. Additionally, we believed that one entrée into the world of academic research is knowing adults who themselves are researchers – either close relationships with adults (not faculty advisors) or family members. Not knowing adults who are researchers may prevent talented students from considering a research career. From the initial applications, we knew 7 of the 10 REU participants did not know adults who were “researchers”. Combined, the program was successful recruiting underrepresented groups. Participants all completed the Collaborative Institutional Training Initiative (CITI) Responsible Conduct of Research Program and attended at least one professional development session on ethical considerations in research and professional practice. Additionally, there were no incidents that would constitute an ethical violation. Program Objectives 2, 3, 5, and 6 were evaluated using data described below.

3.2 Computer & Information Science & Engineering survey

Paired-sample t-tests were used to examine change that occurred after participating in the 10-week REU program on 8 constructs that were measured during the first and final weeks. Mean values for these constructs and for the 3 constructs that were only measured during the final week are found in Table 2. Analyses revealed significant change on Attitudes ($t(9) = 3.70, p = .005, d = 1.04$), Grit ($t(9) = 3.42, p = .008, d = 1.21$), and Research Skills ($t(9) = 2.40, p = .040, d = .71$). All other comparisons were not statistically significant.

Overall, participants’ attitudes toward their project discipline are favorable. However, the results suggest that their favorable attitudes decreased from the beginning to the end of the program.

Two questions that exemplify the *attitudes* construct are 1) I hope that my future career will require the use of [my REU project discipline] concepts, and 2) The challenge of solving problems using [my REU project discipline] appeals to me. An illustration of the shift in attitude was the question that asked 3) Developing (my REU project discipline) skills will be important to my career goals. At the beginning of the program, all students replied somewhat agree to strongly agree (100%). At the end of the program, 80% of the students replied somewhat agree to strongly agree and the other 20% were neutral.

Although a decrease in attitudes toward the project discipline is not a positive outcome, it is not unprecedented. One hypothesis about why this would occur is that participants had unrealistic ideas about the time, effort, or duties of high-level research in math, engineering, and/or biological science disciplines. After the intensive 10-week program, participants' views are more closely aligned with the reality of research work in the area.

Participants' Grit also decreased from the beginning to the end of the program. As an example, one question asked participants to report about whether new ideas and projects distract them from previous ideas and projects. At the beginning, most participants were not distracted easily by new projects, reporting that being distracted was "not much like me". At the end of the 10-week period, there was a shift in responses so that most students reported being distracted by new projects was "mostly like me". Similarly, there was a shift in responses toward losing interest in certain ideas or projects at the end of the REU program.

There are potentially several reasons why the decrease in *grit* occurred. For example, student reflections referred to the disparity between research activities encountered at the REU site and classroom activities. For example, one student wrote "[REU] learning is a lot harder than learning I've done in class. In class we are given specific instructions for our projects and are taught most of the information we will need to complete the project. The research project I'm working on now involves advanced statistical techniques I didn't know existed until I got here. This is a great learning experience, but it's really stressful because I'm expected to come up with a result when I'm still trying to figure out what's going on."

Cataloging the differences between research and classroom learning are beyond the scope of this paper but there is one more difference that we would point out. Research involves discovery of new knowledge, whereas classroom learning traditionally involves learning what is already known. In research, one must work independently to seek out solutions to research problems and there is a chance for failure. Historically, students who are accepted and participate in REU sites are high achieving. Given the different nature of the activities, many of the students may have experienced more frustration during their REU experience than they had experienced in the classroom.

Importantly, REU participants' self-reported Research Skills increased from the beginning to the end of the program. There were 25 questions that probed their knowledge of, or ability to complete various aspects of research from research proposal write-up to analyzing data with statistics. Participants reported gains in 20

areas with the largest gains in research poster design, conference participation, understanding the nature of the job of a researcher, and ethics in scientific research.

Objectively, students' accomplishments based on the work completed during the 10-week REU program exemplified their increased ability to conduct impactful research. For example, after the REU program, one student continued HPC research under the supervision of an affiliated faculty member and participated in the "Apprentice" level of the XSEDE EMPOWER (Expert Mentoring Producing Opportunities for Work, Education, and Research) program under the project "A Study for Implementation of Parallel Algorithms for High Fidelity Mixtures of Biofilm and Solvent" during the spring 2019 semester. The student will also participate in the Student Research Participation (SRP) program at Argonne National Laboratory and the PETSc group in summer 2019.

Another student worked on parallel implementation of AC optimal power flow using PETSc DMNetwork which is a newer parallel data structure developed by Argonne National Laboratory. Based on his implementation, the affiliated faculty member's research team published a referred conference paper in the IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC 2019). The paper was awarded "Best paper in the track". The student's contribution during the REU program was vital for the research success of the faculty member's team and further collaboration with the PETSc group of Argonne National Laboratory.

Table 2. CISE Survey pre- and post-test results

CONSTRUCT	Pre-Test	Post-Test	Sig.
	5/31/2018	8/2/2018	
Self-Efficacy	3.55	3.84	$p = .100$
Intentions toward Graduate School	3.78	3.63	$p = .452$
Attitudes	4.42	3.85	$p = .005$
Help Seeking (coping)	4.27	3.91	$p = .095$
Grit	3.39	2.91	$p = .008$
Research Skills	2.81	3.32	$p = .040$
Scientific Leadership	4.19	4.08	$p = .493$
Scientific Identity	3.00	3.20	$p = .477$
Faculty Mentor	--	4.28	--
Graduate Student Mentor	--	3.70	--
Overall Program	--	4.27	--

Note. Pre-Test and Post-Test mean scores for all constructs including significance values. All items are on a 5-point Likert type rating scale (higher numbers indicate greater endorsement of construct).

3.3 The Torrance Tests of Creative Thinking

Innovations that drive progress in solving global challenges require people to think creatively. Creativity requires knowledge. Over the 10-week experience, participants gained a significant amount of knowledge of many topics, including HPC. We examined REU site participants' creativity before and after the 10-week program. Our hypothesis was that new knowledge may increase participants' creativity on objective measures. The results were mixed.

Overall, participants exhibited above average creativity that increased from pre- to post-test on the verbal assessment, $t(9) = .997, p = .345$, and decreased from pre- to post-test on the figural assessment $t(9) = .673, p = .518$. Although mean values increased and decreased on the verbal and figural assessments respectively, neither comparison reached statistical significance and thus in the interest of caution we can conclude there was no change. Individual components of creativity, such as fluency (i.e. the number of ideas) and originality (statistical rarity) can be separated to more closely inspect the effect of the program on creativity. In summary, participants gained a significant amount of knowledge on several topics; however, their increased knowledge of research topics and techniques did not translate to increased creativity on the TTCT. Future work in the area could emphasize concrete strategies to improve ideation, originality and elaboration on the TTCT but could also include measures of creativity that are specific to HPC or the project discipline domain.

Participants' written responses also included some reference to creativity in their daily lab work indicating gains in HPC or project discipline creativity. When given a prompt to consider the role of a creative engineer to explore and scrutinize existing data and to generate novel solutions, most participants provided an example of "daily" creativity. One student wrote about her challenges creating a new design for a heat exchanger. She wrote "My project involves creating a new heat exchanger design. I used the literature available to see what kinds of new things people were trying or wanted to test. I then combined two of the things – trapezoid tubes and wavy channels to see if it would create an even better design. I am not sure if the design will be better or if it is too much to try at once."

The participant's description of her research on heat exchangers includes references to several of the requisite skills to generate novel solutions. She explored existing knowledge, found applications to her research problem, analyzed existing designs, and synthesized elements of multiple designs to generate a novel heat exchanger. Finally, she referred to a need for evaluating the product after implementation.

3.4 Interviews and Participants' Written Responses

REU site participants wrote and submitted daily logs and reflections regularly. The first author also conducted unstructured interviews. The amount of data collected from participants, in their own words, was extensive. Some themes that emerged and will be explored include preparation and understanding of project goals, scientific communication, and gender equity.

One theme that emerged was that students did not feel prepared for what they would be asked to do during the REU. For example, REU students were asked to begin work immediately on software packages that were completely new to them. For example, students were asked to learn R, SolidWorks, Simcenter Star-CCM+, Fortran

and more. This is somewhat to be expected – students were entering a new lab with new research procedures and new areas of research.

For context, when students were selected for the SDSU REU they received a notification that named the faculty member they would work with. It is our understanding that for REU sites, the National Science Foundation discourages extensive contact with the students prior to their REU experience because some students will not be able to work ahead. If some students prepared and others were not able to prepare, for various reasons, it could create inequities.

One suggestion to mitigate students' feelings of being ill-prepared would be to communicate with the students as they begin the program that there will be a normal "onboarding" period when they will be learning about what is expected.

The REU programs emphasis on scientific communication was noted in most responses. Being able to communicate research objectives, progress, techniques, and other important details to interested parties is an important skill to develop. Many participants were nervous about the scientific communication exercises during the ten-weeks. In all, students completed an "elevator speech", a 5-slide presentation and a longer, more complete presentation of their accomplishments near the end of the program, they met with the South Dakota State University president, and presented a poster presentation at an off-site conference in Pierre, SD. These communication activities were a strong positive of the program, students reported the structured communication activities as very beneficial to their development in science communication.

Another student referred to some gender inequity that she experienced at her home institution and perhaps at the REU site. She was hesitant to give specific examples, but she did suggest that her accomplishments were not recognized as often as her male-counterparts and/or that male-counterparts had "taken credit" for her accomplishments in lab meetings. As one example, the student had found a journal article that addressed a topic the lab was working on. She shared the journal article with another male REU student who then took credit for finding the article in a later lab meeting. A participating faculty mentor reported hearing some of these concerns as well. The participating faculty mentor conveyed a message very similar to the message conveyed by the external evaluator. That is, students should take steps to discuss their role in lab accomplishments and to take credit where credit was due. In most instances, responses were complimentary of the program, but they also conveyed a certain amount of challenge and in some cases frustration about the learning curve associated with a transition to high-level research.

4 CONCLUSIONS

The authors believe this project expanded participation in STEM by recruiting a diverse pool of first- and second-year students from smaller colleges, tribal colleges, and community colleges. Participants gained a broad perspective on the theory and applications of HPC and the responsible conduct of research. This, in conjunction with professional development activities (technical communication, leadership skills, etc.), will provide students with a competitive edge in the science and engineering workforce.

We asked REU participants to reflect on their experiences in addition to logging their activities. Certainly, reflection is an

important component of progress in the discovery process. Not only should students have documented when something succeeded or failed, they should have reflected and evaluated successes and failures. Reflection takes time, instead of moving on to the next phase or step of the project, it requires time to think about the process from a broader perspective. Some students were more committed to this process than others. If participating faculty and staff members were able to encourage and model a reflective practice it may increase “buy-in” from REU participants.

A complete evaluation of our holistic summer undergraduate research program in high performance computing would necessarily include long-term tracking of the students who participated. At present, the authors have received communications or have knowledge about a small number of the participants. Of those participants, we know that research continues to be an important element of their education. Former participants are involved in research examining recurrent neural networks, soft robotics, electroactive polymer actuators, and others. Future REU sites should include longer-term assessment strategies that participants are aware of and should expect. In fact, the CISE survey could be administered at more than two time-points. Furthermore, a systematic reporting of tangible products (e.g., conference proceedings and publications) that may have come from the projects begun during the REU or overall would be beneficial for comparison with other non-REU, but similar aptitude level students.

In summary, through experience, training, and mentoring, participants in the REU program at SDSU gained the skills, aptitude, and inclination to pursue further study and careers in STEM areas.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1559978. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] Lathrop, S., 2016. A Call to Action to Prepare the High Performance Computing Workforce, *Computing in Science & Engineering* 18, 80; <https://doi.org/10.1109/MCSE.2016.101>.
- [2] National Academies of Sciences, Engineering and Medicine. 2016. *Future Directions for NSF Advanced Computing Infrastructure to Support US Science and Engineering in 2017-2020*. Washington, D.C.: The National Academies Press, doi:10.17226/21886.
- [3] Newhall, T., Danner, A., and Webb, K., 2017. Pervasive parallel and distributed computing in a Liberal Arts College Curriculum, *J. Parallel Distrib. Computing*, Vol 105, pp. 53-62. <http://dx.doi.org/10.1016/j.jpdc.2017.01.005>.
- [4] C. Manners and S. Gent, “Data Analysis and Visualization for High Fidelity CFD Models for Medical Device Realization,” in *ASME 2017 International Mechanical Engineering Congress and Exposition*, Tampa, Florida, USA, November 3-9, 2017, vol. V003T04A063, 2017.
- [5] A. Werner, K. Duwadi, N. Stegmeier, T. M. Hansen, and J.-H. Kimn, Parallel implementation of ac optimal power flow and time constrained optimal power flow using high performance computing, in *IEEE Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2019, p. 6 pp. (Best paper Award).
- [6] Salas, P., Fennell, A., and Swaminathan, P., 2018. Grape Berry Ripening Signature Gene Expression Analysis, in *South Dakota EPSCoR Undergraduate Research Symposium*, July 26, 2018.

- [7] Rorrer, A., 2016. An evaluation capacity building toolkit for principal investigators of undergraduate research experiences: A demonstration of transforming theory into practice. *Evaluation and Program Planning*, 55, 103-11.
- [8] Torrance E.P., 1974. *Torrance tests of creative thinking: Norms and technical manual*. Bensenville, IL: Scholastic Test Services.