The Use of HFOSS Projects in the Grace Hopper Celebration of Women in Computing Open Source Day

Dr. Cam Macdonell, MacEwan University

Cam Macdonell is an Assistant Professor of Computer Science at MacEwan University.

Heidi J.C. Ellis, Western New England University

Heidi Ellis is a Professor in the Computer Science and Information Technology department at Western New England University. Dr. Ellis has a long-time interest in software engineering education and has been interested in student participation in Humanitarian Free and Open Source Software (HFOSS) since 2006. She has received multiple NSF grants supporting research into student participation in HFOSS.

Darci Burdge, Nassau Community College

Darci Burdge is Assistant Chair and Professor of Computer Science at Nassau Community College. She has worked to increase students understanding of software development and the impact it can have on society. She is especially interested in broadening the perspective of the introductory Computer Science student beyond the programming concepts typically taught in these courses. She uses HFOSS projects as a means to providing real-world experience and finds that students are motivated, showing increased participation in classroom discussion especially among women. She is Co-PI on an NSF-funded project to assist faculty who are interested in involving students in HFOSS projects.

Dr. Lori Postner, Nassau Community College

Lori Postner is a Professor of Computer Science at Nassau Community College in Garden City, NY. She is Co-PI on an NSF grant focused on helping faculty to incorporate student participation in HFOSS (Humanitarian Free and Open Source Software) projects in their CS curriculum.

Dr. Gregory W Hislop, Drexel University (Computing and Informatics)

Gregory Hislop is a Professor and Senior Associate Dean in the College of Computing and Informatics at Drexel University. His scholarly interests span computing education research, information technology for teaching and learning, and software engineering. Prior to coming to Drexel, Dr. Hislop spent 18 years working in government and industry, where his efforts included software development and support, technology planning and evaluation, and development and delivery of technical education.
The Use of HFOSS Projects in the Open Source Day at the Grace Hopper Celebration of Women in Computing

Abstract
The Grace Hopper Celebration of Women in Computing is the world’s largest meeting for women in technology. The Open Source Day, a hackathon using humanitarian free and open source software (HFOSS) projects, has been held in conjunction with the conference for the past seven years. Open Source Day provides an unusual opportunity to survey a large number of women as they contribute to HFOSS projects. More typically, the small number of women computing students makes it difficult to measure the impact of HFOSS participation on women’s attitude toward computing. Although geared toward students, the Open Source Day attracts women of all skill levels and backgrounds. This paper reports on a pre/post survey designed to understand women’s opinions of HFOSS participation and how those opinions may change when working on a humanitarian project. The results demonstrate HFOSS has wide-spread appeal among women and that women have interest in humanitarian applications of computing. Our survey participants represented 28 different countries, across age groups, work experience and programming background.

1. Introduction
The Grace Hopper Celebration of Women in Computing was founded in 1994 and is the world’s largest technical conference for women in computing. More than 15,000 people from 87 countries attended the 2016 conference with an even greater number attending in 2017 [5]. Many students attend the conference to network, learn and find community, and a large number of technical companies attend to learn how to build inclusive cultures as well as to recruit technical talent.

Since 2010, the Hopper conference has hosted an Open Source Day [6]. Originally titled "Open Source Code-a-thon for Humanity", the day provides women of varying skill levels with a day-long experience in contributing to humanitarian free and open source software (HFOSS). Projects include Systers, Women’s P2P Network, the Center for Open Science, and humanitarian projects supported by Red Hat and Mozilla. Open Source Day is attended primarily by students, but also attracts professors and professionals.

The set of projects available for attendees is primarily determined by the organizations that volunteer to participate. Projects are advertised prior to the conference and attendees pre-register for the project to which they would like to contribute. Attendees work in groups throughout the day and, at the end of the day, groups present their work to the day’s participants as well as the sponsoring organizations. Few groups manage to make any significant contribution to a project, but the day does provide some in-depth introduction to the project, often under the guidance of someone with considerable project knowledge.

Computer Science (CS) programs have recently experienced a surge in enrollments surpassing the growth of the late 1990’s [9]. There has also been an increase in the number of non-majors taking CS courses. Despite there being more students in CS courses, the number of women graduating from CS programs was still only 15% in 2015 [9]. Since HFOSS projects are built
with the intention to better the human condition, they often exemplify characteristics that 
research has shown attracts women to computing such as helping others and working in diverse 
groups [2, 10]. There is some evidence that working on HFOSS projects is appealing to women 
[13, 14]. The 2016 Open Source Day offered a unique opportunity to study the impact that 
contributing to an HFOSS project has on participant motivation with an audience comprised 
almost entirely of women. As the majority of Open Source Day participants are students, it could 
be expected that the study would provide insight into the impact of a single day event on female 
student motivation to study computing.

Involvement in HFOSS provides contributors with a range of learning experiences from 
technical skills to professional skills. HFOSS projects make project artifacts visible and the 
community interacts in public forums. This project transparency and community accessibility 
make HFOSS projects fertile learning environments [7, 8, 12]. Recent research has shown that 
involvement in HFOSS within a class can positively impact student attitude towards computing, 
major selection and career plans, technical and professional knowledge, software engineering 
knowledge, and more [15, 16]. In addition, research has shown that women demonstrate a 
significantly greater understanding than men of how to plan, develop and maintain an HFOSS 
project after having participated in an HFOSS community [13]. The research presented in this 
paper explores the impact of HFOSS on a significantly sized group of women in computing, 
largely made up of students.

2. Method

The study was designed to understand if a one-day exposure to HFOSS motivates females with 
regard to computing majors or careers. A single set of research hypotheses that include a null and 
directional hypothesis are investigated in this study:

- H10: Short-term participation in an HFOSS project has no impact on motivation to 
  study/work in computing.

- H1a: Short-term participation in an HFOSS project has positive impact on motivation to 
  study/work in computing.

The study used a pre/post anonymous survey that asked participants to provide a unique 
identifier to allow for matching pre- and post-surveys while maintaining anonymity. The goal of 
the survey was to obtain participant opinions about computing before and after participating in 
the Open Source Day. The survey was constructed to be relatively short to improve response 
rate. The survey contained demographic information (country of origin, age, gender, ethnicity, 
major) and information about participant experience with computing (self-rated programming 
experience and computing work experience). Five-point Likert scale responses (Strongly 
Disagree, Disagree, Neutral, Agree, Strongly Agree) were used to gauge participant opinion of 
the experience. Options for “Don’t know” and “Not applicable” were also provided. In addition, 
the pre- and post-surveys each contained a single open-ended question about what the participant 
expected to learn (pre-survey) and the most important thing that was learned (post-survey). Table 
1 shows the Likert items from the pre-experience survey. Similar items using the past tense were 
used for the post-experience survey.

The survey did not specifically ask if the participant was a student, faculty or computing 
professional, an oversight that will be corrected in future surveys. The age ranges (18-19, 20-21, 
22-23, >= 24) were intended to help distinguish where the participant might be in their
education. A work experience question was intended to identify students who participated in internships as part of their education or as summer work.

3. Results

The original goal of the study was to look at students, however the participant population included faculty and computing professionals as well. Since the survey did not contain a question to differentiate among students, faculty and computing professionals, a combination of age and years of computing work experience was used to get some idea of the likely representation of students within the population. Participants who were 24 or older with more than 2 years of computing work experience were identified as most likely not students.

The Open Source Day had a total of 164 attendees and 41 mentors. A total of 129 subjects participated in the study by completing the pre-experience, post-experience or both surveys, representing a majority of the people who attended the Open Source Day. Two of the participants were excluded as they did not complete enough of the survey to analyze. Of the remaining 127 participants, there were 112 completed pre-experience surveys and 66 completed post-experience surveys. Based upon the unique identifier, there were 51 matched pre/post surveys. It is possible there were more participants that completed both the pre- and post- surveys but the surveys lacked an identifier that could be used to match the surveys.

Using the criteria above, 81 (72%) of the pre-experience surveys and 49 (74%) of the post-experience surveys were likely students. 34 (67%) of the matched surveys were also likely students. The sections that follow summarize the background of the participants, their opinions of the experience, their expectations and post-experience comments.

| H1: I think learning about humanitarian open source projects will give me a better appreciation for the usefulness of computing. |
| H2: I expect that working on a humanitarian open source project will increase my interest in computing. |
| H3: Participation in a humanitarian project may cause me to consider taking more computing courses. |
| H4: I expect that working on a humanitarian open source project will increase my interest in computing as a career. |
| H5: I want to work on a humanitarian open source project because I want to help the people who would benefit from the software. |
| H6: Working on a humanitarian open source project will increase my confidence in my computing ability. |

Table 1: Likert Statements
3.1 Demographics and Background

The subject population consisted of 111 women and 1 man based upon the self-identification question in the survey. 15 participants chose not to answer the question. Participants came from 28 different countries including Indonesia, Iran, Kenya, Pakistan, New Zealand, and more. The largest number of participants were from the U.S. (46) followed by India (34), China (11) and Bangladesh (4). All remaining countries had only one or two participants. The distribution of participants based on continent is shown in Table 2. It is interesting to note that the largest number of participants, almost half, came from Asia.

The survey used four ranks (18-19, 20-21, 22-23, >=24) to capture the age of participants and Table 3 shows the age distribution of participants. It is interesting to note that a large number of attendees were 24 or older. It was also interesting to note that 27 of the 35 participants who identified as being from India were over 24.

Table 4 shows the distribution of ethnicity of the participants. Given the large number of participants from the continent of Asia in Table 2, it is not surprising that the majority of participants indicated that they were of Asian or Pacific Island descent.

Given our survey was focused on students, attendees were ask to give their major area of study. Not unexpectedly 84 participants indicated that they either had or were working on a Computer Science degree. Other majors with more than a single entry included Electrical Engineering, Computer Science & Engineering, Computer Engineering, Information Technology, and Math. While the majority of participants were in some sort of technical discipline, it was interesting to note that there were also attendees from Psychology, Economics, Environmental Science and Languages.

Figure 1 Shows the distribution of self-reported programming ability of the participants. The figure shows a relatively normal distribution with a slight emphasis on higher programming ability. Figure 2 shows the years of computing work experience of participants. Unsurprisingly, the majority of participants had two or fewer years of experience as one would expect from students. Thirty-seven of the participants left this question blank which may mean they had no work experience. Perhaps more interestingly, 12 participants had more than 10 years of computing work experience. Comparing Figure 1 and Figure 2, a natural expectation would be that these two graphs would be somewhat similar as one could assume that as experience increased, so would programming ability. Another interesting aspect is that although there were 56 participants with two or fewer years of experience (and another 37 who probably had none), only 30 participants rated their programming ability as low (1 or 2). On the other end of the scale, even though there were 12 people with 10 or more years of work experience, only seven people rated their programming ability as a 5 (highest value) and only three of those high programmers had 10+ years of experience. One reason for the difference between years of experience and self-reported programming ability may be that women have been shown to be significantly more likely to underestimate rather than overestimate their abilities [1].
### Table 2: Continent of Participants

<table>
<thead>
<tr>
<th>Continent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>60</td>
</tr>
<tr>
<td>Africa</td>
<td>2</td>
</tr>
<tr>
<td>North America</td>
<td>48</td>
</tr>
<tr>
<td>South America</td>
<td>4</td>
</tr>
<tr>
<td>Europe</td>
<td>8</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1</td>
</tr>
<tr>
<td>Unanswered</td>
<td>4</td>
</tr>
</tbody>
</table>

### Table 3: Age of Participants

<table>
<thead>
<tr>
<th>Age</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-19</td>
<td>10</td>
</tr>
<tr>
<td>20-21</td>
<td>33</td>
</tr>
<tr>
<td>22-23</td>
<td>10</td>
</tr>
<tr>
<td>=&gt;24</td>
<td>69</td>
</tr>
<tr>
<td>Unanswered</td>
<td>5</td>
</tr>
</tbody>
</table>

### Table 4: Ethnicity of Participants

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian or Pacific Islander</td>
<td>77</td>
</tr>
<tr>
<td>Black, not Hispanic</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9</td>
</tr>
<tr>
<td>White, not Hispanic</td>
<td>33</td>
</tr>
<tr>
<td>Unanswered</td>
<td>6</td>
</tr>
</tbody>
</table>

### Table 5: Pre/Post-Experience Likert Item Comparison

<table>
<thead>
<tr>
<th>Item</th>
<th>Pre (N=112)</th>
<th>Post (N=66)</th>
<th>Diff</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>4.50</td>
<td>4.54</td>
<td>0.043</td>
<td>0.763</td>
</tr>
<tr>
<td>H2</td>
<td>4.28</td>
<td>4.42</td>
<td>0.147</td>
<td>0.353</td>
</tr>
<tr>
<td>H3</td>
<td>3.75</td>
<td>4.19</td>
<td>0.443</td>
<td>0.020</td>
</tr>
<tr>
<td>H4</td>
<td>4.11</td>
<td>4.27</td>
<td>0.162</td>
<td>0.376</td>
</tr>
<tr>
<td>H5</td>
<td>4.47</td>
<td>4.50</td>
<td>0.027</td>
<td>0.848</td>
</tr>
<tr>
<td>H6</td>
<td>4.18</td>
<td>4.23</td>
<td>0.053</td>
<td>0.749</td>
</tr>
</tbody>
</table>

### Figure 1: Programming Ability

![Programming Ability](image1.png)

### Figure 2: Years’ Experience

![Years’ Experience](image2.png)

### 3.2 Opinion Item Results

In order to analyze the opinion items, the values from “Strongly Disagree” to “Strongly Agree” were assigned integer values from one and five. The average of all responses to the pre- and post-experience opinion items were above neutral as shown in Table 5. Using a 2-tailed t-test, with alpha = 0.10, one item, “H3: Participation in a humanitarian project may cause me to consider taking more computing courses.” showed a statistically significant difference between the pre- and post-test responses.

**Opinion results breakdown by ethnic group** - A further examination of the responses broken down by ethnicity shows a similar pattern. Figure 3 shows a graph of the average responses to the pre-experience survey by ethnicity. It is interesting to note that the averages for “White, not Hispanic” were higher overall than “Asian or Pacific Islander” which were higher overall than
“Hispanic”. Further research into why this may be so would be of interest as the responses to the pre-experience survey provide some insight into the expectations of participants. While the highest average responses overall were from the “Black, not Hispanic” group, there were only two data points in this group. Figure 4 shows the post-survey opinion responses. Note that the average responses for “White, not Hispanic”, “Asian or Pacific Islander”, and “Hispanic” became more closely grouped after the workshop.

The data by ethnic group were analyzed for significant changes using a two-tailed t-test. As noted earlier and shown in Table 5, only question H3 saw a significant change from the pre-experience survey to post-experience survey across all respondents. Table 6 shows the average pre/post experience responses based on ethnicity. (Note that the two data items for the “Black, not Hispanic” group are not represented in Table 6 due to the small number of participants in this category.) The analysis by ethnicity shows a similar pattern to the analysis for all participants. For each ethnicity, using an alpha value of 0.10, a single question shows significant difference, although the question with significant difference varies by ethnicity.

Another interesting aspect is that the significant change in opinion of “White” respondents (H4 “interest in career”) is a negative change, whereas the significant change for the other ethnic groups were positive. This result could indicate that the experience in Open Source Day did not fulfill the participants’ expectations. Or the convergence of the post-experience survey may simply reflect that the women had a better understanding of HFOSS by the end of the day and so response became more similar across ethnicities.

Opinion results breakdown by age - The opinion responses were also analyzed by age categories. Significant positive change was found in age categories “20-21” and “over 24” on H3, “consider taking more courses”. This mirrors the significant change for the total set of respondents. Sample sizes in the other age categories were much smaller and no significant differences were detected.

<table>
<thead>
<tr>
<th>Item</th>
<th>White Pre</th>
<th>White Post</th>
<th>P</th>
<th>Hispanic Pre</th>
<th>Hispanic Post</th>
<th>P</th>
<th>Asia/Pacific Pre</th>
<th>Asia/Pacific Post</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>4.77</td>
<td>4.53</td>
<td>0.318</td>
<td>4.14</td>
<td>4.80</td>
<td>0.36</td>
<td>4.39</td>
<td>4.53</td>
<td>0.483</td>
</tr>
<tr>
<td>H2</td>
<td>4.68</td>
<td>4.32</td>
<td>0.163</td>
<td>3.57</td>
<td>4.80</td>
<td>0.090*</td>
<td>4.14</td>
<td>4.39</td>
<td>0.294</td>
</tr>
<tr>
<td>H3</td>
<td>4.07</td>
<td>4.18</td>
<td>0.758</td>
<td>3.33</td>
<td>4.25</td>
<td>0.396</td>
<td>3.60</td>
<td>4.17</td>
<td>0.037*</td>
</tr>
<tr>
<td>H4</td>
<td>4.70</td>
<td>4.18</td>
<td>0.043*</td>
<td>3.40</td>
<td>4.50</td>
<td>0.287</td>
<td>3.89</td>
<td>4.29</td>
<td>0.146</td>
</tr>
<tr>
<td>H5</td>
<td>4.62</td>
<td>4.60</td>
<td>0.911</td>
<td>4.43</td>
<td>4.67</td>
<td>0.721</td>
<td>4.39</td>
<td>4.38</td>
<td>0.961</td>
</tr>
<tr>
<td>H6</td>
<td>4.57</td>
<td>4.50</td>
<td>0.765</td>
<td>3.80</td>
<td>4.20</td>
<td>0.614</td>
<td>4.00</td>
<td>4.09</td>
<td>0.697</td>
</tr>
</tbody>
</table>

Table 6: Pre/Post-Experience by Ethnicity (* Significant difference for alpha = 0.1)
Pre-Experience Open-ended Observations - The pre-experience survey contained the question “What do you expect to learn by working on a humanitarian open source project?” Seventy-four participants provided an answer to this question. In examining the responses, answers fell into general categories.

A common theme in the pre-survey comments was related to helping society in some way. Comments included:

- Make the world a better place.
- ...how it can help those in need.
- I would like to get more experience and do something meaningful at the same time.
- ...something useful that will help those in need.
- How to better apply my coding to help people.
- How to use my technical skills for good.

Out of 74 respondents, 24 (32%) used words like “help”, “society”, “good”, “change”, or “give”, demonstrating a motivation related to contributing to societal good.

Numerous participants mentioned an interest in learning new technologies or skills, enhancing their current skill set or wanting to understand more about open source in general. Some of the common areas of interest included:

- Web dev/dev on Windows
- How open source works
- Want to gain familiarity with open source workflow
- ...improve my programming skills and communication skills
- Skill up, broaden skills, gain FOSS experience
- Improve coding skills (by coding with others)
These responses are consistent with reasons why someone might participate in a hackathon day. Out of the 74 responses, 27 (37%) used words like “skill”, “learn”, “improve”, and “gain” in their responses. It should be noted that some of the projects in the hackathon were not strictly humanitarian in nature, but the participants used the project for a humanitarian application.

There is a similarity in the respondent proportion who mention contributing to societal good as compared to learning new skills. The advantage of HFOSS applications is that students achieve both goals at the same time.

As the focus of Open Source Day is open source software, it is natural that 21 respondents mentioned “open source” specifically as a motivator for attending this event. Comments on expectations included:

- How to participate in OS projects in the future.
- I want to learn about how control is maintained in big OS projects.
- Want to gain familiarity with open source workflow.

While open source projects may offer the ability to develop skills and experience large-scale software development, the humanitarian appeal adds a further motivation for students to attend.

Post-Experience Open-ended Observations - The post-experience survey contained the question "What was the most important thing that you learned from working on a humanitarian open source project?" Thirty-six participants provided an answer to this question. In examining the responses, answers fell into several general categories.

Given that HFOSS projects are deployed to solve real-world scenarios, their code can be significantly more complex than projects typically found in the classroom. Five respondents made a comment related to the difficulty of setting up the project development environment or the importance of clear documentation:

- There were a lot of technical difficulties we found so we had to find alternative methods to work.
- Set up and start up seems to be as difficult as I previously believed unfortunately.
- That it’s really difficult to get/keep working environments.
- Challenges of handling code that might not be well documented and that may contain bugs.
- To make sure you have good documentation on any installation/set up steps.

In addition to the complexity of the development environment, the sheer number of lines of code in HFOSS projects can make for a steep learning curve. However, one of the important concepts of teaching with large-code bases is that it is not necessary that one understands the entire code base. Some comments reflect the respondents understanding that continuous learning and being “productively lost” are part of working with HFOSS:

- That we are all learning
- That most people, including professors and professionals are as lost as me sometimes
- If you don’t know a lot you can still do a lot
A third observation refers to comments that touch on the enjoyment of working with others and collaborating. Most software is written by teams, and developing team skills is important for producing well-prepared CS graduates. In addition, open source contributions are not solely limited to code contributions and the comments below reflect this reality:

- That there are many ways to get involved.
- Teamwork :-)
- Collaboration makes for better projects
- That women are very collaborative and really enjoy working cooperatively
- Collaboratively/breaking up tasks and asking for help is important.

The open-ended answers of the matched pre- and post-survey were examined to see if there were any interesting changes in participant expectations prior to the experience and their post-experience opinion. A common thread was teamwork/collaboration. Of the 25 matched free response answers, six participants mentioned teamwork in the post-survey, three of whom also mentioned it in the pre-survey. There were two participants who cited working in groups on the pre-survey but not on the post-survey.

4. Discussion

The analysis of the opinion items as well as the open-ended responses provides some significant evidence to support the positive hypothesis H1a: Short-term participation in an HFOSS project has positive impact on motivation to study/work in computing.

The demographics of the women who chose to attend indicate that HFOSS appeals to women across the range of ethnicities and ages represented by Open Source Day participants. The positive opinion responses indicate that HFOSS participation holds potential to motivate women to pursue and continue computing careers. Overall, study results showed more positive responses in comparing the post-experience surveys to the pre-survey. This response is similar to previous studies on the impact of HFOSS experiences on student opinion of computing [14, 15].

Research shows that women and minorities are attracted to majors and careers that have broad impact on their communities and world [3, 4, 10, 11]. The results of the open-ended questions demonstrate that many Open Source Day attendees chose to participate so they could understand the “...importance of computing to daily issues that affect the society”. Without being prompted, many participants indicated that helping others was important to them and that they were looking for a way to use their skills to help others. Many schools have a service learning component in their curriculum and HFOSS participation can be an excellent way to satisfy this requirement. In addition, research has suggested that HFOSS provides many learning opportunities [15]. The data collected indicates that people opted to participate in Open Source Day because they expected to strengthen their technical skills. The post-survey responses reinforced this belief as participants cited what they had learned through their participation. Overall the survey results indicate that participants believe that HFOSS gave them the opportunity to both learn skills and to help others. This finding is important for education as it aligns with the characteristics that should increase the number of women and underrepresented minorities in the field [2, 10].
Finally, it is important to remember the context of this study. The women attending the Open Source Day are mostly undergraduate students who are already interested in computing, and many are already in a computer major. The appeal of the study is that this event provided an opportunity to survey a large group of undergraduate women interested in computing. At the same time, this study does not provide and reliable insight as to how an experience like this might affect attitudes of women who have no serious computing interest. Similarly, this study provides no reliable insight on impact of an experience like this on K-12 students. Both of those populations would be of interest in future work.

5. Conclusion and Future Work

Results of this study show that involvement in HFOSS has broad appeal to women in computing across a variety of demographic factors and that HFOSS participation can reinforce their interest in studying computing. The ability to help society in some way, to "do good", was a common theme among participants. In addition, participants were motivated by the opportunities to learn new skills and technology found in HFOSS projects.

Future work should include investigation of the impact of more substantive participation in HFOSS projects. Treatment effect is difficult to measure in this situation in part because the treatment is very limited.

Acknowledgments

The authors would like to thank the Grace Hopper Convention Open Source Day organizers for their help and support in conducting this research. This material is based on work supported by the National Science Foundation under Grant Nos. DUE-1525039, DUE-1524877, and DUE-1524898. 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 (NSF).

References


