Student Perceptions of Their Abilities and Learning Environment in Large Introductory Computer Programming Courses – One Year Later

Dr. Laura K Alford, University of Michigan

Laura K. Alford is a Lecturer and Research Investigator at the University of Michigan.

Prof. Valeria Bertacco, University of Michigan

Valeria Bertacco is Arthur F. Thurnau Professor of Electrical Engineering and Computer Science and Associate Dean for Physical Sciences and Engineering in the Rackham Graduate School at the University of Michigan. She is also an Adjunct Professor of Computer Engineering at the Addis Ababa Institute of Technology. Her research interests are in the area of computer design, with emphasis on specialized architecture solutions and design viability, in particular reliability, validation and hardware-security assurance. She joined the University of Michigan in 2003, after working with the Advanced Technology Group of Synopsys, which she joined via the acquisition of Systems Science Inc. Valeria has been a member of the Executive Committee of the Design Automation Conference since 2013, serving in the role of Technical Program Chair for 2017. The conference is the top forum in hardware and embedded systems design, attracting over 7,000 attendees yearly. She has also served as Associate Editor of the IEEE Transactions on Computer Aided Design. She is the author of three books on correctness in computer design. She received her M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 1998 and 2003; and a Computer Engineering degree ("Dottore in Ingegneria") summa cum laude from the University of Padova, Italy in 1995. Valeria is the recipient of the IEEE CEDA Early Career Award, NSF CAREER award, the Air Force Office of Scientific Research’s Young Investigator award and the IBM Faculty Award. From the University of Michigan she received the Vulcans Education Excellence Award, the Herbert Kopf Service Excellence Award, the Sarah Goddard Power Award for contribution to the betterment of women and the Rackham Faculty Recognition Award. Valeria is an ACM Distinguished Scientist and an IEEE Fellow.
Student Perceptions of Their Abilities and Learning Environment in Large Introductory Computer Programming Courses – One Year Later

Abstract

Over the past 30 years, women completing computer science and computer engineering undergraduate degrees have been a minority compared to their male counterparts. Three obstacles to gender diversity in computer science and computer engineering are: stereotyped traits, perceived abilities, and learning environment. Identifying implicit bias as a component of these obstacles, we implemented a series of activities designed to lessen the impact of implicit bias on our students in large-enrollment introductory computer programming courses. One element of assessing the success of our program is to use entry and exit surveys to gauge the change in students’ perceptions of their abilities and learning environment. In particular, we are interested in the difference between men’s and women’s perceptions of their abilities and the learning environments in these courses.

The initial findings of these entry and exit surveys found that while there are differences between men’s and women’s responses, the differences were not as great as we had feared. However, due to the relatively large number of responses (1200+) it is possible that even a small difference in, for example, student agreement/disagreement with “I believe that other students in computer programming courses will be welcoming of me” could have a disproportionately large effect on the number of women deciding to major in computer science/computer engineering.

After improving the survey process based on recommendations from the initial study, we embarked on a 5 year program to gather data and assess the gender differences in two sequential large programming courses. Our overarching research question is: Do women and men show a statistically significant difference in their perceptions of their abilities and learning environment as measured by self-efficacy, intimidation by programming, and feelings of inclusion?

This paper will present the first set of entry and exit survey results (Fall 2017) analyzed using mixed model ANOVA for repeated measures of questions on self-efficacy, intimidation by programming, and feelings of inclusion. Statistically significant results include: women have lower self-efficacy than men in both courses, and women are more intimidated by programming than men in the second programming course. Although we cannot reject the null hypothesis for any of our three hypotheses regarding these questions, we can still glean useful insight from this data set.
Introduction and Motivation

Over the past 30 years, women completing computer science and computer engineering undergraduate degrees have been a minority compared to their male counterparts.\(^1\) The reasons for this gender gap are complex,\(^2\) but they can be generalized into the problems of recruitment and retention.

A previous study by the authors\(^3\) looked at ways to remove or minimize the impact of three obstacles to gender diversity in the computer science and computer engineering undergraduate programs: stereotyped traits, perceived abilities, and learning environment. Interventions implemented in the first and second programming courses include:

- balanced teaching staff in terms of gender and race (visual representation is critical)
- staff training on implicit bias, stereotype threat, etc.
- various student activities related to implicit bias
- personalized messaging via an electronic coaching system

The obstacles of stereotyped traits, perceived abilities, and adverse learning environment exist for many different social groups. These introductory programming courses are some of the largest engineering classes offered at this institution; as such, they have a wide sphere of influence on the student body. An improved understanding of student experiences in these classes will provide guidance on creating and sustaining a welcoming environment for all students.

In our previous research, we asked students for their perceptions of the impact of stereotypes, preconceived notions of ability, and learning environment on their experiences in their programming courses. The difference in responses between women and men were not as large as we had feared at the beginning of the study, and this was encouraging. This may indicate that our efforts to combat implicit bias were partially successful, although since we do not have pre-implicit bias activity data, we cannot be certain. However, the relatively large number of responses (1200+) means it is possible that even a small difference in responses between men and women may be significant.

With this previous study appearing to indicate that the programming courses are not horrible in terms of gender disparity, we turn now towards incremental improvement and longer term trends. This paper addresses three action items from the previous study:

- Improvement #1: Continue to administer the surveys and use results to guide future course development and other possible interventions.
  
  Action: We have IRB approval for a 5 year program to assess the success of this program on combating implicit bias.

- Improvement #2: Ask demographic information at the end of the entry survey to avoid tainting data with preconceived notions of what the answers should be (stereotype threat).
  
  Action: We moved all demographic-based questions to the end of the survey.
Improvement #3: Analyze the revised surveys more rigorously to determine metrics that are statistically significant.

*Action:* We have processed the new, revised surveys using mixed model ANOVA.

Recruiting and retaining women and underrepresented minorities requires term-by-term assessment of students’ perceptions of the courses. In particular, we chose to focus on indicators for self-efficacy, intimidation by programming, and inclusion. This paper details the entry and exit survey questions used to gather data for these indicators, the analysis results, and our comments on the comparisons between women and men at the beginning of the term and at end of the term.

**Methods**

We conducted entry and exit surveys to investigate the change in students’ perceptions of their abilities and the learning environment in their large introductory programming course. Our hope was that women’s perceptions of their abilities and learning environment would show a positive change from the entry survey to the exit survey if our implicit bias activities were effective. Entry and exit surveys already existed for each of the introductory programming courses participating in this study; therefore, we added the questions we needed to the existing surveys. The participants and specific questions added are detailed in the next sections.

**Participants**
The entry and exit surveys were administered to students in two different large-enrollment computer programming courses at a large, midwestern research university. Course 1 is required of all engineering students and is taken in the first year of study. The students in Course 1 were offered a small amount of extra credit to complete these surveys. Course 2 is the next programming class that students take; it consists of mostly first- and second-year engineering students. Course 2 is required for some majors (computer science, computer engineering, electrical engineering, etc.) but is not required for many other majors, although many non computer/electrical engineering students do take Course 2 in order to become more proficient at programming. The students in Course 2 were required to complete these surveys, which were attached to the existing entry and exit surveys, for course credit. All students were offered the opportunity to opt in or out to the research program without affecting the credit earned in the class. Only students 18 years or older are included in this research.

**Design**
The surveys were designed to be parallel as much as possible to allow for direct comparison of students’ perceptions at the beginning of the course and at the end of the course.

**Indicator 1: Self-Efficacy.** The definition of success differs widely from student to student, but it primarily revolves around grades. For example, an open-ended question on the entry survey was, “Briefly describe what being SUCCESSFUL in *Course 1* means to you.” A word cloud constructed from the students’ answers is shown in Fig. 1, and grades are the most commonly referenced definition of success.
Some studies find that women have lower self-efficacy in STEM fields\(^4\) while others show almost no gender gap.\(^5\) In computer programming, women may have lower computer self-efficacy initially, but they may have a faster rate of increase over the semester than their male counterparts.\(^6\) Self-efficacy can affect the grade a student earns;\(^7\) if the students earns or exceeds the grade they want, then they feel more successful, leading to an increase in self-efficacy. In computer programming classes, this may lead to the student being more likely to declare computer science/engineering as a major.\(^8\) Our measure of change in self-efficacy is the pair of questions, “How confident are you in your ability to be successful in this course?” and “Do you think you were successful in this course?” If our efforts are effective, then we would see an improvement in women’s self-efficacy over the course of the term.

Hypothesis 1: Women have lower self-efficacy in our programming courses, as compared to their male peers, but show improvement between the start of term and the end of term.

**Indicator 2: Intimidation.** Intimidation is the summation of many things, and women on average are at a disadvantage when it comes to when they were first introduced to computers;\(^9\) what experiences, role models, and encouragement they had in K-12,\(^10\) and previous programming experience.\(^11,12,13\) These experiences, taken together, can lead to an increased feeling of intimidation when women are required to take a programming course. As we do not have any control over our students’ previous programming experiences, we would like to at least slightly lessen how intimidated students feel between the beginning of the term and the end of the term. If our efforts are effective, then we would see a decrease in intimidation over the course of the term.

Hypothesis 2: Women are more intimidated by programming in our programming courses, as compared to their male peers, but are less intimidated by the end of term.
**Indicator 3: Inclusion.** A major aspect of college is peer interaction. Stereotypes certainly play a role in whether a woman will feel welcome and valued in a field, and hence choose to major in that field. Peer encouragement and respect are also important when students choose a major and whether they stay in that major. Therefore, tracking students’ feeling of welcome over time will give us an idea of how the inclusiveness of the undergraduate CS program may be changing. If our efforts are effective, then we would see an increase in student’s perception of being welcome over the course of the term, with women reporting a larger increase over the course of the term.

Hypothesis 3: Women feel less welcome in our programming courses, as compared to their male peers, but feel more welcome by the end of term.

Table 1 shows the questions used to gather data for these three indicators. The questions as included in the surveys are in Appendix: Sample Survey Questions on page 13.

Table 1: Survey questions used for assessment. Sample questions are included in the Appendix.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Entry Survey</th>
<th>Exit Survey</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>How confident are you in your ability to be successful in this course?</td>
<td>Do you think you were successful in this course?</td>
<td>linear scale</td>
</tr>
<tr>
<td>Intimidation</td>
<td>I find computer programming intimidating.</td>
<td>I find computer programming intimidating.</td>
<td>linear scale</td>
</tr>
<tr>
<td>Inclusion</td>
<td>I believe that other students in computer programming courses will be welcoming of me.</td>
<td>I believe that other students in computer programming courses will be welcoming of me.</td>
<td>linear scale</td>
</tr>
</tbody>
</table>

**Analysis**

We have two within-subjects repeated measures: the start of term and the end of term. There are two between-subjects: women and men. There are three indicators (self-efficacy, intimidation by programming, and inclusion) and two courses; the survey questions for these indicators are in Table 1. The same survey questions are used for both courses.

After the entry and exit survey data were administered, we first removed any student who was not over 18 and all students who declined to have their answers used for research. Then, the surveys were matched up by student identifier to find the students that took both the entry and the exit surveys. After the matching was done, we removed the student identifiers. From this set of data, we selected those students who self-reported as either men or women. The data set was first graphed as histograms to get a visual sense of the distribution of the answers, then we used SPSS to analyze each data set using a general linear model for repeated measures (mixed model ANOVA).
Results

A summary of the number of responses and response rates is shown in Table 2.

The distribution of responses for each pair of questions are shown in Figs. 2–4. The visual distributions are histograms plotted horizontally with men’s responses on the left and women’s responses on the right. The y axis is labeled with the responses (linear scale or multiple choice); the x axis is the percentage of men or women who selected that response.

The descriptive statistics and marginal mean plots from the mixed model ANOVA are presented in Figs. 5–6. More detailed results are located in the Appendix on page 13.

Table 2: Summary of responses for those students who took both the entry and the exit surveys in the Fall 2017 term (September-December 2017).

<table>
<thead>
<tr>
<th>Course</th>
<th>Enrollment</th>
<th>Total # Respondents</th>
<th>Response Percentage</th>
<th>Women Respondents</th>
<th>Men Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>778</td>
<td>232</td>
<td>29.8%</td>
<td>80</td>
<td>152</td>
</tr>
<tr>
<td>2</td>
<td>921</td>
<td>535</td>
<td>58.1%</td>
<td>171</td>
<td>364</td>
</tr>
<tr>
<td>Total</td>
<td>1699</td>
<td>767</td>
<td>45.1%</td>
<td>432</td>
<td>775</td>
</tr>
</tbody>
</table>

Figure 2: Distribution of responses to the questions, “How confident are you in your ability to be successful in this course?” (Entry Survey) and, “Do you think you were successful in this course?” (Exit Survey)
Figure 3: Distribution of responses to the statement, “I find computer programming intimidating.” Course 1 appears to be divisive for men: male students who were neutral in the beginning of the course grow to either actively like, or actively dislike, programming.

Figure 4: Distribution of responses to the statement, “I believe that other students in computer programming courses will be welcoming of me.”
(a) Self-Efficacy. Between-subjects (gender) is statistically significant.

(b) Intimidation by Programming.

(c) Inclusion.

Figure 5: Summary of mixed model ANOVA results for Course 1. Under “gender”, 0=Men and 1=Women. See Appendix for more details.
(a) Self-Efficacy. Between-subjects (gender) is statistically significant.

(b) Intimidation by Programming. Between-subjects (gender) is statistically significant.

(c) Inclusion.

Figure 6: Summary of mixed model ANOVA results for Course 2. Note the reduced range on the vertical axis for the plot of “Inclusion.” See Appendix for more details.
Discussion

Here we offer some discussion of general trends and patterns of this data. We will focus on results that were statistically significant. See the Appendix on page 13 for the significance of within- and between-subjects effects for each course’s three indicators.

Self-Efficacy
Visually inspecting the survey results for self-efficacy in Figs. 2, 5a, and 6a, we see that overall men report more self-efficacy than women in both courses and that this difference is statistically significant. This agrees with the large amount of previous research that also finds these differences, and supports the first part of our Hypothesis 1. In Course 1, women show an increase in self-efficacy over the term, but this increase is not statistically significant, therefore we cannot reject the null hypothesis for Hypothesis 1. In Course 2, students of both genders studied show a slight decrease in self-efficacy, but again this is not statistically significant. The histogram view in Fig. 2 shows the asymmetry in men’s and women’s reported self-efficacy (reflected in the marginal means), and also shows how the increase in Course 1’s women’s self-efficacy comes primarily from those women who were neutral at the beginning of the course feeling more successful at the end of the class.

Intimidation
Visually inspecting the survey results for intimidation by programming in Figs. 3, 5b, and 6b, we see that overall women are more intimidated by programming in both courses, although this difference is statistically significant only for Course 2. Therefore, we cannot reject the null hypothesis for Hypothesis 2. Interestingly, we see that men at the start of Course 2 are, on average, about the same as where they left off at the end of Course 1, but women are on average more intimidated (not statistically significant). We wonder if this means the women in Course 2 are truly more intimidated by programming or whether they are just more realistic about starting a more difficult programming course! The estimated means for men in Course 1 change very little from the beginning of the term to the end of the term. The histogram view in Fig. 3 gives insight into this apparent non-change – Course 1 appears to be divisive for men in terms of being intimidated by programming. Men who were neutral about programming in the beginning of the term have split, with some being less intimidated by programming and some more intimidated. Women do not report this particular shift.

Inclusion
Visually inspecting the survey results for inclusion in Figs. 4, 5c, and 6c, we do not see much difference between men and women, between the time periods, or between the courses. No effects are statistically significant; therefore, we cannot reject the null hypothesis for Hypothesis 3. Overall, students seem to feel fairly welcome, though the few that do not may well be students that are underrepresented in computer science/engineering. Neither course causes a large decrease in inclusion, which is encouraging.
Conclusions and Recommendations

In this study, we presented the statistical analysis for three indicators of student success (self-efficacy, intimidation by programming, and inclusion) in two sequential large computer programming courses. There are two within-subjects repeated measures: the start of term and the end of term. There are two between-subjects: women and men.

Statistically significant results include: women have lower self-efficacy than men in both courses, and women are more intimidated by programming than men in the second programming course. Although we cannot reject the null hypothesis for any of our three hypotheses, we can still glean useful insight from this data set.

Instructors of large computer programming courses need to continue to support women to overcome obstacles of lower self-efficacy and intimidation by programming. Reinforcing that women “do this” by having female instructors and female teaching assistants is critical, but those women need to be careful not to reinforce existing negative stereotypes of women who program. Diversity in all interpretations of the word is key.

For our own efforts, we plan to continue using these surveys to build more knowledge about our students and inform how we can support them best. For self-efficacy, we plan to calculate the change in response (e.g. from neutral (3) to agree (4)) and do a mixed mode ANOVA on these results to see whether there are large or small increases/decreases in self-efficacy. For intimidation by programming, we plan to identify those students (men and women) who have taken both Course 1 and Course 2 and analyze their responses at the four times (beginning and end of each term) to gain a clearer understanding of how students progress through this programming sequence. For inclusion, we plan to move away from gender differences and begin to look at the inclusion of underrepresented minority students.

Acknowledgement

This research was supported by the Computing CARES program at the University of Michigan, which is supported by a 2nd Place NCWIT NEXT Award. We appreciate the recognition and support of the National Center for Women and Information Technology! This research has been determined exempt from human subjects control under exemption #1 of the 45 CFR 46.101.(b) by the U-M Institutional Research Board (HUM00135376).

References


Appendix: Sample Survey Questions

Appendix: Mixed Mode ANOVA Results

The results of the mixed mode ANOVA for all cases is included here. Recall that Course 1 is a required intro to programming course for first year engineering students. For Course 2, the next programming course in the programming sequence. Students are primarily first and second year engineering students. Course 2 is required for some majors, but not required for others (though many non-computer-science/engineering students take this course).
Figure 7: Results for Course 1: Self-Efficacy. Between-subjects (gender) is statistically significant.
Figure 8: Results for Course 1: Intimidation. Under “gender”, 0=men and 1=women. No statistical significance.
Figure 9: Results for Course 1: Inclusion. Under “gender”, 0=men and 1=women. No statistical significance.
Figure 10: Results for Course 2: Self-Efficacy. Between-subjects (gender) is statistically significant.
Figure 11: Results for Course 2: Intimidation. Between-subjects (gender) is statistically significant.
Figure 12: Results for Course 2: Inclusion. Note the reduced range on the vertical axis. No statistical significance.