



Student Sense of Community Through an Introductory Computer Programming Course Sequence

Dr. Laura K Alford, University of Michigan

Laura K. Alford is a Lecturer and Research Investigator at the University of Michigan. She researches ways to use data-informed analysis of students' performance and perceptions of classroom environment to support DEI-based curricula improvements.

Dr. Amir Kamil, University of Michigan

Dr. Andrew DeOrio, University of Michigan

awdeorio@umich.edu contact Andrew DeOrio is a teaching faculty member at the University of Michigan and a consultant for web and machine learning projects. His research interests are in ensuring the correctness of computer systems, including medical and IOT devices and digital hardware, as well as engineering education. In addition to teaching software and hardware courses, he teaches Creative Process and works with students on technology-driven creative projects. His teaching has been recognized with the Provost's Teaching Innovation Prize, and he has twice been named Professor of the Year by the students in his department.

Student Sense of Community Through an Introductory Computer Programming Course Sequence

Abstract

An inclusive community is an important factor in growing and maintaining a diverse student body. In this paper, we measure changes in the sense of community among students as they progress through an introductory computing sequence.

A five year program is aimed at increasing the enrollment and graduation rates of women and underrepresented minority (URM) students in computer science and engineering at a competitive public research institution. Three obstacles to diversity in computer science and computer engineering have been identified: stereotyped traits, perceived abilities, and learning environment. Identifying implicit bias and imposter syndrome as components of these obstacles, we include a series of class activities designed to lessen the impact of implicit bias and imposter syndrome 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.

Previously, we investigated the difference between men's and women's perceptions of their abilities and the learning environments in introductory computing courses. We found a statistically significant association between gender and perception of self-efficacy, but not in how those perceptions change over the course of the term. We then investigated the differences between underrepresented minority (URM) students' and non-URM students' perceptions of their abilities and the learning environments in introductory computing courses. We found no statistically significant differences between URM and non-URM students; however, both groups showed slight decreases in feelings of inclusion for the courses included in the study. To better understand these previous results, we need a more nuanced investigation of individual paths over time through the introductory computer programming sequence.

Our research question is: How do students' sense of community change over time through an introductory computer programming sequence?

Our data set is comprised of entry and exit survey data for five semesters of a three course introductory computing sequence, which yielded 259 complete responses. We measured changes in survey responses from individual students as they progressed from CS1 through CS2 to CS3. The survey data was analyzed using a mixed-effects linear model for repeated measures of questions on the student's sense of community in their undergraduate studies up to the point of when they took the survey.

The data show that students in all groups report generally positive feelings for every survey questions, and that mean values are fairly consistent across groups. However, we did observe several statistically significant effects, indicating a change in sense of community and self-efficacy. Overall, students report a small but significant decrease over time in response to questions related to self-efficacy as they progress through the program. Women in particular show a stronger negative effect compared to men. URM students show small to medium negative effects for questions related to climate and inclusion as they make progress through the course sequence. These results indicate that we require further interventions to improve the sense of community among our students.

Introduction and Motivation

The number of AP CS exams offered and passed still lags other AP exams, and the number of women and URM students who pass the AP CS exams is even smaller. [1] URM women are particularly under-represented in these exams. AP exams are not the deciding factor in high school students' decisions about what to study in college, but they are an indicator of how broadly acceptable different fields are to different social groups. [2] Visualizing the AP exam data,[3] we see that many exams have more women taking the exams than men (the English and other language exams, in particular), but the CS exams are among the most male-skewed exams.

Belongingness and a Sense of Community

We know that belongingness is important in engineering contexts[4] and that self-identifying as “an engineer” increases students' sense of belonging[5]. However, a binary evaluation of belongingness persists: students either feel that they have the skills and belong or they do not[6], which can cause problems because we know that it is important to have a growth mindset rather than a fixed mindset[7]. White students often have a stronger sense of belonging than black students[8], and a lack of role models for women contributes to a lower sense of belonging for female engineering students[9]. Additionally, computer science is not always perceived as “friendly to society” which can cause students who want to help society to feel less like they belong in computer science and engineering[10].

Students' perceptions and beliefs about their engineering identity are just as important as their math and science skills[11]. Broader examples of using computing to solve engineering problems will allow students to each find something of interest in their classes, positively impacting their sense of belonging and community[12]. In general, we need to de-emphasize the more famous applications of computer science (game design, app development), and instead emphasize the myriad ways that computing helps society[10, 13].

Our Introductory Computer Programming Sequence

Our computer science and engineering program wishes to increase sense of community among students taking the three-course introductory computer programming sequence, with a special focus on women and underrepresented minority (URM) students as these groups may be at a disadvantage when beginning the sequence [14, 15, 16, 17, 18]. We started a 5 year program to

coordinate interventions across courses and assess the effect of these interventions. Interventions include[19]:

- balanced teaching staff in terms of gender and race (visual representation is critical)
- staff training on implicit bias, imposter syndrome, and stereotype threat
- student activities related to implicit bias and imposter syndrome
- in-class examples featuring diverse characters and situations
- personalized messaging via an electronic coaching system
- video dispelling myths about the third course being a “weeder course”

Previous research generated by this 5 year program include a gender analysis[20] and a race analysis[21] of students’ responses to the statement “I believe that other students in computer programming courses will be welcoming of me” on entry and exit surveys for the first two courses of the the three-course sequence. According to these analyses, women felt welcome, but men felt slightly more welcome compared to women, and URM students felt equally welcome compared to non-URM students.

These were positive findings, but we were concerned that the results were limited by assessing this single question for a measure of inclusion. Therefore, we proposed to conduct a more nuanced investigation of students’ feelings of inclusion by analyzing a set of thirteen questions aimed at capturing students’ overall sense of community in these large computer programming courses.

This paper details the entry and exit survey questions used to gather data for this study, the analysis and results, and our comments on how students’ sense of community changes as they travel through this three-course introductory programming sequence.

Methods

We conducted entry and exit surveys to investigate the evolution of students’ sense of community in their large introductory programming courses. Our hope was that women and URM students’ sense of community would increase throughout the three course sequence as each course builds upon the previous course(s). Entry and exit surveys already exist for each of the introductory programming courses participating in this study; the participants and specific questions used for this study are detailed in the next sections.

Participants and Courses

The entry and exit surveys were administered to students in four different large-enrollment computer programming courses at a large, public research institution.

Course 1a is a CS1-type course and is required of all engineering students and is taken in the first year of study; the students in Course 1a were offered a small amount of extra credit to complete these surveys. Course 1b is also a CS1-type course and is available for all non-engineering students to take; the students in Course 1b were required to complete these surveys for course

credit. Because both Course 1a and 1b feed into Course 2, the responses from both courses are combined as “CS1” in this study.

Course 2 is the next programming class that students take; it consists of first- and second-year 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 take Course 2 in order to become more proficient at programming. The students in Course 2 were required to complete these surveys for course credit. In this study, these responses are referred to as “CS2”.

Course 3 “Data Structures and Algorithms”, follows Course 2 in the programming sequence. Course 3 consists of mostly second year students who intend to major or minor in Computer Science. Students learn to write complex programs optimized for speed and efficiency. The students in Course 3 were also required to complete the surveys for course credit. In this study, these responses are referred to as “CS3”.

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 who opted in to the research program are included in this research.

Design

The surveys were designed to be parallel to allow for direct comparison of students’ ratings from one moment in time to another. Each survey included 13 questions intended to capture a student’s overall sense of community, derived from the Basic Psychological Need Satisfaction and Frustration Scale [22]:

Thinking about your experiences in your undergraduate studies so far, please indicate how true each statement is for you on a scale of 1 (Not at all true) to 5 (Extremely true). Each question also included a “Not Applicable” option.

1. I feel a sense of choice and freedom in what I undertake
2. I feel capable at what I do
3. I really like the people I interact with
4. I feel confident that I can do things well
5. I feel that my decisions reflect what I really want
6. I feel included in the groups that I want to belong to
7. I feel competent to achieve my goals
8. I get along with people I come into contact with
9. I feel my choices express who I really am
10. I feel I am doing what really interests me
11. People are generally pretty friendly towards me

12. I feel I can successfully complete difficult tasks

13. I feel optimistic about my career prospects after I complete my education

Our research question is: How do students' sense of community change over time through an introductory computer programming sequence?

Hypothesis: Women and under-represented minority students feel less a part of the community, as compared to men, at the start of CS1, but they feel more a part of the community by the end of CS3 than they do at the start of CS1.

We hypothesize that the factors discussed in the Introduction and Motivation section lead to women and URM students feeling less a part of the community than non-URM men as they enter CS1, when measured by the full set of questions above. However, we expect that our interventions in CS1, CS2, and CS3 improve the climate such that women and URM feel more integrated into the community as they proceed through the three-course sequence. In this work, we analyze the responses to the survey questions above to determine whether the interventions are enough to accomplish this.

Note that we are not evaluating the interventions directly – we do not have survey data from before the interventions were in place, nor with a control group that does not receive the interventions. As such, our work is focused on measurement and analysis, which can then help guide what further interventions are necessary to improve climate.

Analysis

We have six potential within-subjects repeated measures: the start of term and the end of term for each of three courses in the introductory computer programming sequence. There are six potential between-subjects measures: URM students/non-URM students and women/men/non-cisgender. There are 13 survey questions; the same survey questions are used for all surveys. The survey data in this study is from Fall 2017 – Fall 2019 (inclusive).

Survey Matching. After the entry and exit survey data were administered, each student was assigned a unique identifier using a hash function based on students' institutional IDs. We removed any student who was not over 18 and all students who declined to have their answers used for research. The surveys were matched by unique identifier to find the students who completed both the entry and the exit surveys for each of the three courses: CS1, CS2, and CS3; we retained only those responses for students who completed all six surveys. If a student repeated a course, we retained only their most recent survey response.

Junk Responses. We removed any data set that included one or more sets of "junk responses." We categorize a junk response as an entry survey with the same numerical answer for all questions paired with an exit survey that also has the same numerical answer for all questions; the actual number used may be different between the entry survey and the exit survey. We also categorized a response as junk if the exit survey has more than one "Not Applicable" marked because all questions should be applicable at the end of a course.

Race. We classified students as URM or non-URM students. To classify students as URM vs. non-URM, we started with the definition used by our institution, which in turn relies on the NSF

definition of URM students. Our institution defines URM students as persons who identify as African-American/Black, Hispanic, Hawaiian, or Native American. Therefore, we categorized all students who identified as solely "White" or solely "Asian" as non-URM students and all other students as URM students.

Gender. We classified students as women, men, or non-cisgender. We started with the students' self-reported gender on the surveys. We categorized students who selected "Woman" as women, students who selected "Man" as men, and all other students as non-cisgender.

Visual and Statistical Analysis. A sum score (the sum of the responses to each of the 13 survey questions) was used as an overall indicator of a student's sense of community. Aggregate responses to each survey question, along with the sum score, were plotted to visualize general trends. Responses are categorized by gender and URM status; means and mean +/- standard deviation are plotted for each distinct group in the gender/URM intersection. We used MATLAB to analyze the data set using a mixed-effects linear model to determine the statistical significance of the data set (13 survey questions + sum score).

Results

There were 7,843 total students who took the CS1 courses from Fall 2017 – Fall 2019. There were 6,835 students that completed the entry surveys; this is our starting point as we attempt to track student sense of community across the introductory programming sequence. Table 1 shows the overall response rates for the CS1, CS2, and CS3 courses.

Table 1: Summary of overall response rates across the courses.

| Course | Enrollment | # students who took | | # students who took | |
|--------|------------|---------------------|------------|-----------------------------|------------|
| | | the entry survey | response % | both entry and exit surveys | response % |
| CS1 | 7843 | 6835 | 87.1 | 3847 | 49.1 |
| CS2 | 5195 | 4165 | 80.2 | 2455 | 47.3 |
| CS3 | 2305 | 2101 | 91.1 | 1088 | 47.2 |
| total | 15343 | 13101 | 85.4 | 7390 | 48.2 |

Notes: CS1 includes both Course 1a and Course 1b

CS1/CS2 include data from F17 – F19; CS3 includes data from F18 – F19

After post-processing all survey responses according to the process described in the Analysis section, we had just 261 qualifying sets of responses. A summary of the race and gender breakdown of these 261 sets of responses is shown in Table 2.

There were two non-cisgender students who completed entry and exit surveys for all three courses; this is too few responses for us to draw any conclusions on non-cisgender students at this time. Therefore, we further restricted this study's gender analysis to those students who self-reported as women and men (N = 259).

There were 23 URM students that completed entry and exit surveys for all three courses. Although this sample size is small, we included the effects of URM status in our model.

Table 2: Summary of responses for those students who took *both* the entry and the exit surveys in CS1, CS2, and CS3 courses from Fall 2017 through (and including) Fall 2019 terms. N = 261.

| | NUMBER OF RESPONDENTS | | | PERCENTAGE OF RESPONDENTS | | | |
|----------------|-----------------------|-------|---------------|---------------------------|-------|---------------|-----|
| | Men | Women | Non-Cisgender | Men | Women | Non-Cisgender | |
| Non-URM | 142 | 94 | 2 | Non-URM | 54.4 | 36.0 | 0.8 |
| URM | 12 | 11 | 0 | URM | 4.6 | 4.2 | 0.0 |

Therefore, we have four distinct groups represented in our data set for this study:

1. URM women
2. non-URM women
3. URM men
4. non-URM men

Means and Standard Deviations

The means and standard deviations for the sum scores for each group on the six surveys are listed in Table 3. The maximum possible sum score is 65 (65 = 13 questions × 5 (Extremely True)).

Table 3: Means and standard deviations of the sum scores on each survey for each of the four groups. The maximum possible sum score is 65. (65 = 13 questions × 5 (Extremely True)).

N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

| Survey | | URM women | | non-URM women | | URM men | | non-URM men | |
|--------|-------|-----------|-------|---------------|-------|---------|-------|-------------|-------|
| | | mean | stdev | mean | stdev | mean | stdev | mean | stdev |
| CS1 | Entry | 49.89 | 6.75 | 49.75 | 7.01 | 52.89 | 8.62 | 52.20 | 5.63 |
| | Exit | 50.64 | 3.53 | 50.68 | 7.38 | 51.17 | 11.06 | 53.95 | 6.93 |
| CS2 | Entry | 52.50 | 3.31 | 49.96 | 7.22 | 52.00 | 8.08 | 52.47 | 6.44 |
| | Exit | 52.55 | 7.85 | 48.53 | 7.39 | 51.17 | 8.79 | 52.73 | 7.20 |
| CS3 | Entry | 52.09 | 6.25 | 48.78 | 7.95 | 48.00 | 12.99 | 52.26 | 6.47 |
| | Exit | 48.00 | 6.53 | 47.84 | 8.93 | 47.67 | 8.73 | 51.25 | 8.05 |

Fig. 1 shows the trends of the means of the sum scores. Fig. 2 shows the spread of the sum scores across the surveys using filled bands that represent the mean sum score +/- one standard deviation. Mean and standard deviation plots for all 13 questions are in the Appendix on page 15.

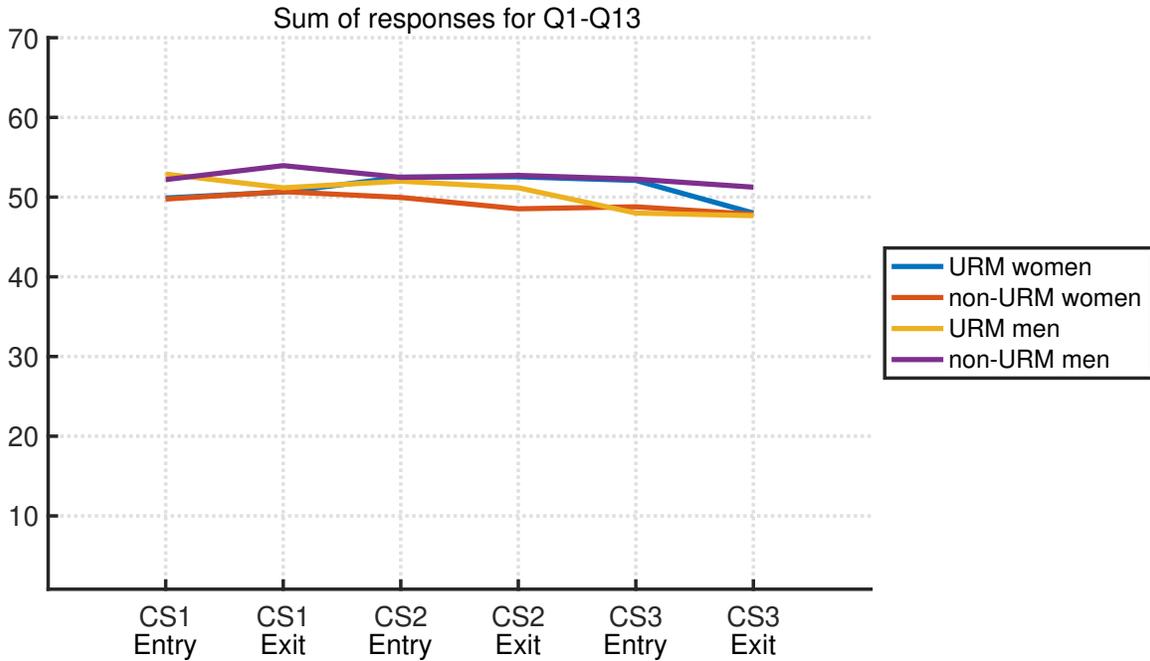


Figure 1: Means of the sum scores (sum of the responses to Q1-Q13).
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

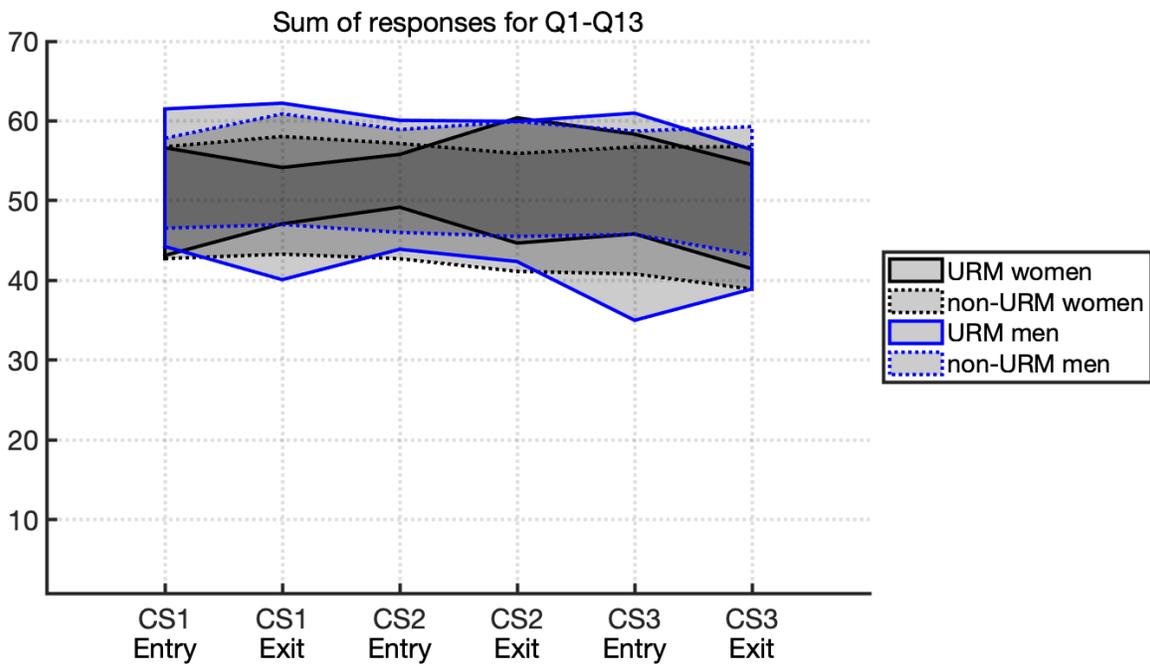


Figure 2: Spread of the sum scores (sum of the responses to Q1-Q13); filled bands are the mean +/- one standard deviation.
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

Statistical Analysis

We restrict our reported results here to those that are statistically significant. Table 4 summarizes the statistically significant results of fitting a linear mixed-effects model to the data set ($p < 0.05$). In the “Name” column, “time” indicates that the results were significant across the different points in the course sequence, and “time:URM” indicates that the combination of time and URM ethnicity was significant.

Table 4: Statistically significant results of fitting a linear mixed-effects model to the data set ($p < 0.05$).

Effects that also meet the $p < 0.01$ criteria are highlighted in blue .

Effects that also meet the $p < 0.001$ criteria are **highlighted in blue and bolded** .

N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

| Question | Name | Estimate | SE | tStat | DF | pValue | Lower | Upper |
|------------|--------------|----------------|---------------|----------------|-------------|---------------|----------------|----------------|
| Q1 | time | -0.0316 | 0.0158 | -1.9980 | 1540 | 0.0459 | -0.0626 | -0.0006 |
| Q2 | time | -0.0426 | 0.0143 | -2.9780 | 1526 | 0.0029 | -0.0707 | -0.0146 |
| Q2 | Women | -0.2634 | 0.0976 | -2.6991 | 1526 | 0.0070 | -0.4548 | -0.0720 |
| Q3 | time | -0.0302 | 0.0142 | -2.1273 | 1539 | 0.0336 | -0.0580 | -0.0024 |
| Q4 | time | -0.0461 | 0.0157 | -2.9357 | 1539 | 0.0034 | -0.0768 | -0.0153 |
| Q4 | Women | -0.3784 | 0.0992 | -3.8142 | 1539 | 0.0001 | -0.5730 | -0.1838 |
| Q5 | Women | -0.3209 | 0.1025 | -3.1308 | 1537 | 0.0018 | -0.5219 | -0.1198 |
| Q6 | time:URM | -0.1243 | 0.0423 | -2.9360 | 1528 | 0.0034 | -0.2073 | -0.0412 |
| Q7 | time | -0.0477 | 0.0144 | -3.3223 | 1539 | 0.0009 | -0.0759 | -0.0196 |
| Q7 | Women | -0.2710 | 0.0913 | -2.9681 | 1539 | 0.0030 | -0.4501 | -0.0919 |
| Q10 | Women | -0.2466 | 0.0988 | -2.4949 | 1530 | 0.0127 | -0.4404 | -0.0527 |
| Q11 | time:URM | -0.0659 | 0.0310 | -2.1245 | 1539 | 0.0338 | -0.1268 | -0.0051 |
| Q12 | time | -0.0466 | 0.0140 | -3.3347 | 1537 | 0.0009 | -0.0741 | -0.0192 |
| Q12 | Women | -0.3512 | 0.0930 | -3.7752 | 1537 | 0.0002 | -0.5337 | -0.1687 |
| Q13 | time | -0.0362 | 0.0157 | -2.3043 | 1532 | 0.0213 | -0.0671 | -0.0054 |
| Q13 | Women | -0.3238 | 0.1078 | -3.0040 | 1532 | 0.0027 | -0.5352 | -0.1124 |
| Sum Q1-Q13 | time | -0.3607 | 0.1234 | -2.9231 | 1501 | 0.0035 | -0.6028 | -0.1187 |
| Sum Q1-Q13 | Women | -2.6631 | 0.8663 | -3.0741 | 1501 | 0.0021 | -4.3623 | -0.9638 |

Discussion

Here we offer some discussion of general trends and patterns of this data based on the visual representation of the responses from the four different groups (URM women, non-URM women, URM men, and non-URM men). In the statistical analysis, we focus on results that were statistically significant.

Overall Trends and Impressions

Visually inspecting the survey results across all 13 questions shown in the Appendix on page 15 shows that overall these students report feeling generally positive about their sense of community.

All mean values are > 3 , indicating that students are generally positive when thinking about their sense of community.

The highest sum score mean (Table 3 or Fig. 1) is the 53.95 mean of the sum scores reported by the non-URM men at the end of CS1. We might expect this group to have the highest ratings at this point as they are “stereotypical engineering students” and could be considered to feel the most comfortable with the overall community as they are in the majority. However, this highest mean sum score, 53.95, is only 83% of the theoretical maximum sum score of 65. One might interpret this as: *even the “most comfortable students”* only feel like roughly 80% part of the community. In general, all groups show mean values that are fairly consistent with each other and across the course sequence.

Looking at the standard deviations as an indicator of the spread of the sum scores, we do see the variation in student responses. URM men show the widest variation, but we do note that the sample size for this group is small ($N = 12$). The darkest band, representing where the four groups overlap, ranges from roughly 44-55 out of the possible maximum score of 65. The “middle score” would be $13 \text{ questions} \times 3 \text{ ratings} = 39$, so the bulk of students do land on the positive side of the range.

We do note a few specific trends of interest from the individual questions:

- *Question 2: I feel capable at what I do.* In Fig. 5, the mean non-URM women’s response is lower than the other groups’ mean response at each time point, although still above the neutral “3” rating.
- *Question 4: I feel confident that I can do things well.* In Fig. 9, the mean non-URM women’s response is lower than the other groups’ mean response at each time point, although still above the neutral “3” rating.
- *Question 6: I feel included in the groups that I want to belong to.* In Fig. 13, we see a decline in URM men reporting that they feel included in the groups that they want to belong to. URM women show increases in feeling that they belong to the groups that they wish to after taking CS1 and CS2, but show a decline after taking CS3.

Statistically Significant Results

All statistically significant effects are negative, indicating a decrease in sense of community for all groups, with the decrease felt more strongly especially for women overall. This unfortunately means we must reject our hypothesis.

Questions 1-13 only have a range of 1-5 on the ratings scale. Therefore, an effect size greater than 0.2 would be considered “large to medium” and an effect size greater than 0.4 is certainly large. The “Estimate” column in Table 4 shows the effect size. Most of the statistically significant effects were for time and women, with more limited effects on time*URM status. We discuss these effects in the next sections, but we note that these measures are for the groups (URM/non-URM/Women/Men) as a whole; an individual student of any group may still be at-risk of feeling excluded from the community depending on their own personal circumstances.

Time. Many questions, along with the sum score, show statistically significant negative effects for time. However, the effect sizes are small at less than 0.05 for each of the questions that show a negative effect. Most of the questions with significant negative effects for time are related to confidence and efficacy: Q2 *I feel capable at what I do* (-0.0426), Q4 *I feel confident that I can do things well* (-0.0461), Q7 *I feel competent to achieve my goals* (-0.0477), Q12 *I feel I can successfully complete difficult tasks* (-0.0466), Q13 *I feel optimistic about my career prospects after I complete my education* (-0.0362).

URM Status. We did not observe a statistically significant effect for URM status alone. However, two questions had a significant effect in combination with time: Q6 *I feel included in the groups that I want to belong to* (-0.1243) and Q11 *People are generally pretty friendly towards me* (-0.0659). While these effect sizes are small to medium, it is notable that these questions are focused on inclusion and climate, and that a significantly negative effect was only observed for URM students as they progress through the program.

Women. The largest, most significant effects on women were on questions 4 and 12. These effects are both negative and stronger (-0.3784 and -0.3512, respectively). Significant negative effects were also observed on questions Q2 (-0.2634), Q7 (-0.271), and Q13 (-0.3238). As with time, these are the questions related to confidence and efficacy. Women also reported significantly lower numbers on questions Q5 *I feel that my decisions reflect what I really want* (-0.3209) and Q10 *I feel I am doing what really interests me* (-0.2466). This suggests a lower degree of satisfaction with their academic program than for other groups.

Limitations

In this study, several factors were beyond our control. Here, we discuss these threats to validity.

Our data set is from a longitudinal study that depends on student participation at the beginning and end of three courses. Student responses are voluntary and the number of individuals whose data is complete is low compared to the number of students enrolled in these large courses. It is possible that there is a self-selection bias among the respondents.

The responses in our data set are on a 1 to 5 integer scale, and thus the granularity of the responses is limited. Furthermore, the sample size is relatively small (259), which limits the effect size that is observable.

Additionally, the students in our study were drawn from a single institution. If the population of students at our competitive public research institution is not representative of students at other institutions, that could limit the ability of our results to generalize.

Finally, there is no control group. Our program consists of a series of activities designed to lessen the impact of implicit bias and imposter syndrome, material that is presented to all students in our courses. Without a control group, we were not able to compare groups of students exposed to the material against those who were not. Additionally, we do not control for instructor or other outside influences.

Conclusions and Recommendations

In this paper, we examined the experience of URM students in introductory computer programming courses at a large, public research institution. In an effort to reduce obstacles to student success, the courses include learning modules designed to lessen the impact of implicit bias and imposter syndrome, staff training on these topics and stereotype threat, diverse teaching assistants, and other activities to promote climate and inclusion. Students completed pre- and post-surveys in three courses, and our data set includes responses from 259 students that completed all six of the surveys.

We presented the statistical analysis for thirteen questions related to students' sense of community across the three courses in our study along with the sum score of the thirteen questions. We analyzed six within-subjects repeated measures: the start of term and the end of term for each of the three courses. We also measured the association between-subjects: URM and non-URM students, as well as cisgendered men and women.

We observed that for all questions and all groups, students reported generally positive feelings with mean values > 3 on a five-point scale. In general, mean values for each question are fairly consistent across different groups and over time. However, we observed several statistically significant negative effects, indicating a decrease in the sense of community for all groups, but with a stronger decrease for women. Questions having to do with self-efficacy show small but statistically significant negative effects over time, and these questions also show stronger negative effects for women. Small to medium negative effects were also observed for URM students over time on questions focused on inclusion and climate.

These findings indicate that our program still has more work to do to support students' sense of community and self-efficacy, and that the current interventions need to be augmented. Women in particular report more negative feelings with respect to self-efficacy, and URM students show a significant decrease in inclusion over time as they progress through the three-course sequence.

We have also identified several new potential research questions for the future:

- Do these findings differ, if at all, when disaggregated by multiple social groups (race/ethnicity + gender + sexual orientation)? This will likely require a true longitudinal study to produce enough data for each of these sub-groups.
- Do these findings differ after controlling for prior experience, performance in each course, or overall GPA?
- Do these findings differ after controlling for instructor race and gender?
- Do these findings differ if we are able to capture the whole (or nearly the whole) student population taking these courses? The data in this work represent only 11% of the students who took CS3, and an even smaller fraction of those who took CS1 or CS2.
- Are there specific activities, such as the exercise on implicit bias or an exercise on values affirmation (which has been shown to decrease URM achievement gap[23]), that have a measurable effect on students' perceptions and grades in these programming courses?

Acknowledgement

This research was supported by the Computing CARES program at the University of Michigan. 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). The authors wish to thank the Consulting for Statistics, Computing, & Analytics Research (CSCAR) Group at the University of Michigan for their advice and guidance on the statistical analysis in this paper.

References

- [1] College Board. AP program participation and performance data 2019, accessed: February 2, 2020. URL <https://research.collegeboard.org/programs/ap/data/participation/ap-2019>.
- [2] Marina Papastergiou. Are computer science and information technology still masculine fields? high school students' perceptions and career choices. *Computers & Education*, 51(2):594 – 608, 2008. ISSN 0360-1315. doi: <https://doi.org/10.1016/j.compedu.2007.06.009>. URL <http://www.sciencedirect.com/science/article/pii/S0360131507000759>.
- [3] Barbara Ericson. Slides on AP CSA and CSP data, accessed: February 2, 2020. URL <https://cs4all.home.blog/2020/01/13/slides-on-ap-csa-and-csp-data/>.
- [4] D. Wilson, D. Spring, and L. Hansen. Psychological sense of community belonging in engineering education. In *2008 38th Annual Frontiers in Education Conference*, pages F3F–21–F3F–24, Oct 2008. doi: 10.1109/FIE.2008.4720650.
- [5] Dina Verdín, Allison Godwin, Adam Kirn, Lisa Benson, and Geoff Potvin. Understanding how engineering identity and belongingness predict grit for first-generation college students. In *2018 CoNECD - The Collaborative Network for Engineering and Computing Diversity Conference*, Crystal City, Virginia, April 2018. ASEE Conferences. URL <https://peer.asee.org/29589>.
- [6] Jacqueline Ann Rohde, Lisa Benson, Geoff Potvin, Adam Kirn, and Allison Godwin. You either have it or you don't: First year engineering students' experiences of belongingness. In *2018 ASEE Annual Conference & Exposition*, Salt Lake City, Utah, June 2018. ASEE Conferences. URL <https://peer.asee.org/31320>.
- [7] Carol S. Dweck. Mind-sets and equitable education. *Principal Leadership*, 10(5):26 – 29, 2010. ISSN 21562113.
- [8] Dong Zhao, Denise Rutledge Simmons P.E., and Meltem Duva. Measuring students' class-level sense of belonging: A social-network-based approach. In *2019 ASEE Annual Conference & Exposition*, Tampa, Florida, June 2019. ASEE Conferences. URL <https://peer.asee.org/33093>.
- [9] Lisa Benson, Candice Bolding, Jennifer Harper Ogle, Catherine McGough, Joseph Murphy, and Rachel Lanning. Engineering students' perceptions of belongingness in civil engineering. In *2019 ASEE Annual Conference & Exposition*, Tampa, Florida, June 2019. ASEE Conferences. URL <https://peer.asee.org/32737>.
- [10] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. Alignment of goals and perceptions of computing predicts students' sense of belonging in computing. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, ICER '19, page 11–19, New York, NY, USA, 2019. Association

for Computing Machinery. ISBN 9781450361859. doi: 10.1145/3291279.3339426. URL <https://doi.org/10.1145/3291279.3339426>.

- [11] Anita D. Patrick, Maura J. Borrego, and Carolyn Conner Seepersad. A combined model for predicting engineering identity in undergraduate students. In *2018 ASEE Annual Conference & Exposition*, Salt Lake City, Utah, June 2018. ASEE Conferences. URL <https://peer.asee.org/29660>.
- [12] Meagan R. Kendall, Nathan Hyungsok Choe, Maya Denton, and Maura J. Borrego. Engineering identity development of hispanic students. In *2018 ASEE Annual Conference & Exposition*, Salt Lake City, Utah, June 2018. ASEE Conferences. URL <https://peer.asee.org/30411>.
- [13] Timothy J. Weston, Wendy M. Dubow, and Alexis Kaminsky. Predicting women's persistence in computer science- and technology-related majors from high school to college. *ACM Trans. Comput. Educ.*, 20(1), September 2019. doi: 10.1145/3343195. URL <https://doi.org/10.1145/3343195>.
- [14] R. A. Ash, J. L. Rosenbloom, L. Coder, and B. Dupont. *Personality characteristics of established IT professionals I: Big Five personality characteristics*. IGI Global, Hershey, PA, 2006. pp. 983-989.
- [15] Nicole Zarrett, Oksana Malanchuk, Pamela Davis-Kean, and Jacquelynne Eccles. Examining the gender gap in it by race: Young adults decisions to pursue an it career. *Women and information technology: Research on underrepresentation*, pages 55–88, 02 2006.
- [16] S. Katz, J. Aronis, D. Allbritton, C. Wilson, and M. L. Soffa. Gender and race in predicting achievement in computer science. *IEEE Technology and Society Magazine*, 22(3):20–27, Fall 2003. ISSN 0278-0097. doi: 10.1109/MTAS.2003.1237468.
- [17] Susan Haller and Sylvia Beyer. Gender differences and intragender differences in computer science students: are female cs majors more similar to male cs majors or female nonmajors? *Journal of Women and Minorities in Science and Engineering*, 12:337–365, 01 2006.
- [18] Jun He and Lee Freeman. Are men more technology-oriented than women? the role of gender on the development of general computer self-efficacy of college students. *Journal of Information Systems Education*, 21:672, 01 2009.
- [19] L. K. Alford, M. L. Dorf, and V. Bertacco. Student perceptions of their abilities and learning environment in large introductory computer programming courses. *Proceedings of the 2017 Annual Conference of the American Society for Engineering Education*, 2017.
- [20] L. K. Alford and V. Bertacco. Student perceptions of their abilities and learning environment in large introductory computer programming courses – one year later. *Proceedings of the 2018 Annual Conference of the American Society for Engineering Education*, 2018. URL <https://peer.asee.org/31012>.
- [21] L. K. Alford and A. DeOrio. Student perceptions of their abilities and learning environment in large introductory computer programming courses – underrepresented minority students. *Proceedings of the 2019 Annual Conference of the American Society for Engineering Education*, 2019. URL <https://peer.asee.org/33297>.
- [22] Beiwen Chen, Maarten Vansteenkiste, Wim Beyers, Liesbet Boone, Edward Deci, Jolene van der Kaap-Deeder, Bart Duriez, Willy Lens, Lennia Matos, Thanasis Mouratidis, Richard Ryan, Kennon Sheldon, Bart Soenens, Stijn Van Petegem, and Joke Verstuyf. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, 39:216–236, 04 2014. doi: 10.1007/s11031-014-9450-1.
- [23] Hannah Jordt, Sarah L. Eddy, Riley Brazil, Ignatius Lau, Chelsea Mann, Sara E. Brownell, Katherine King, and Scott Freeman. Values affirmation intervention reduces achievement gap between underrepresented minority and white students in introductory biology classes. *CBE—Life Sciences Education*, 16(3):ar41, 2017. doi: 10.1187/cbe.16-12-0351. URL <https://doi.org/10.1187/cbe.16-12-0351>. PMID: 28710060.

Appendix: Aggregate Responses

We summarize the responses to all the survey questions in the following graphs, with responses categorized as coming from URM women, non-URM women, URM men, and non-URM men. The first graph for each question tracks the mean numerical value of the responses for each group across the six survey points. The second graph shows the spread of the scores for each group and survey point, using filled bands that represent the mean score \pm one standard deviation.

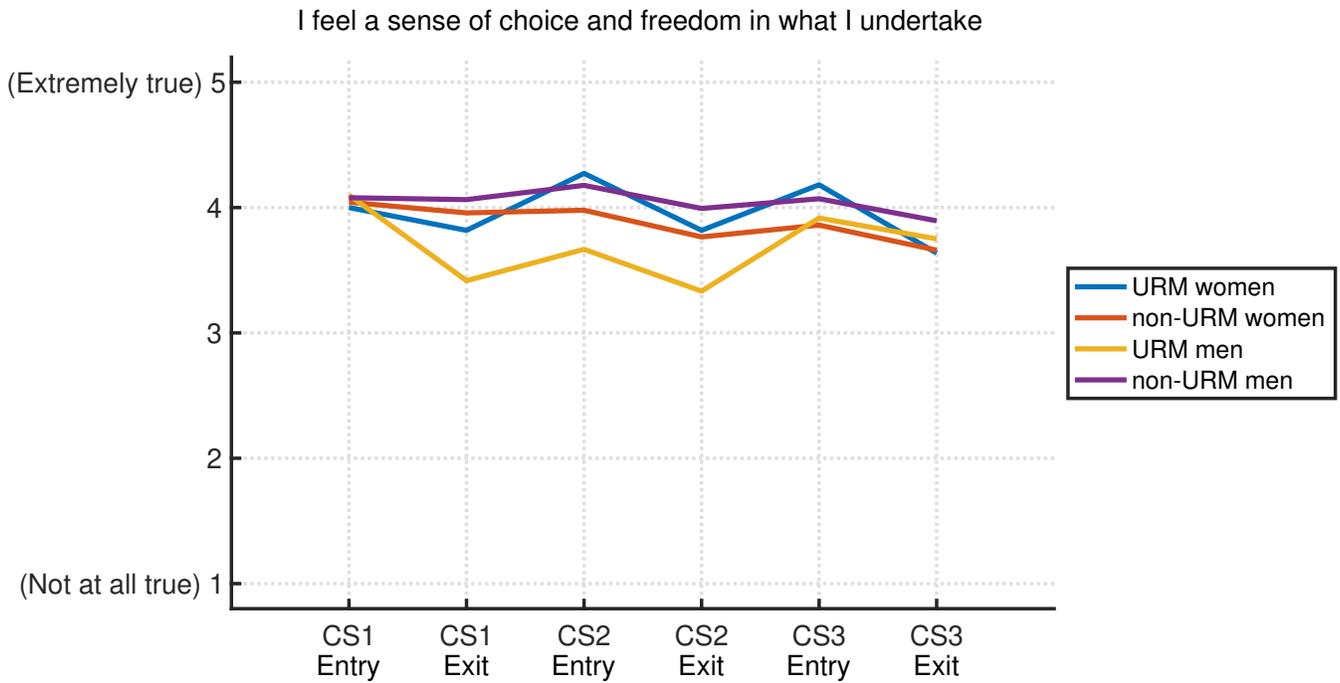


Figure 3: Means of the responses to Question 1.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

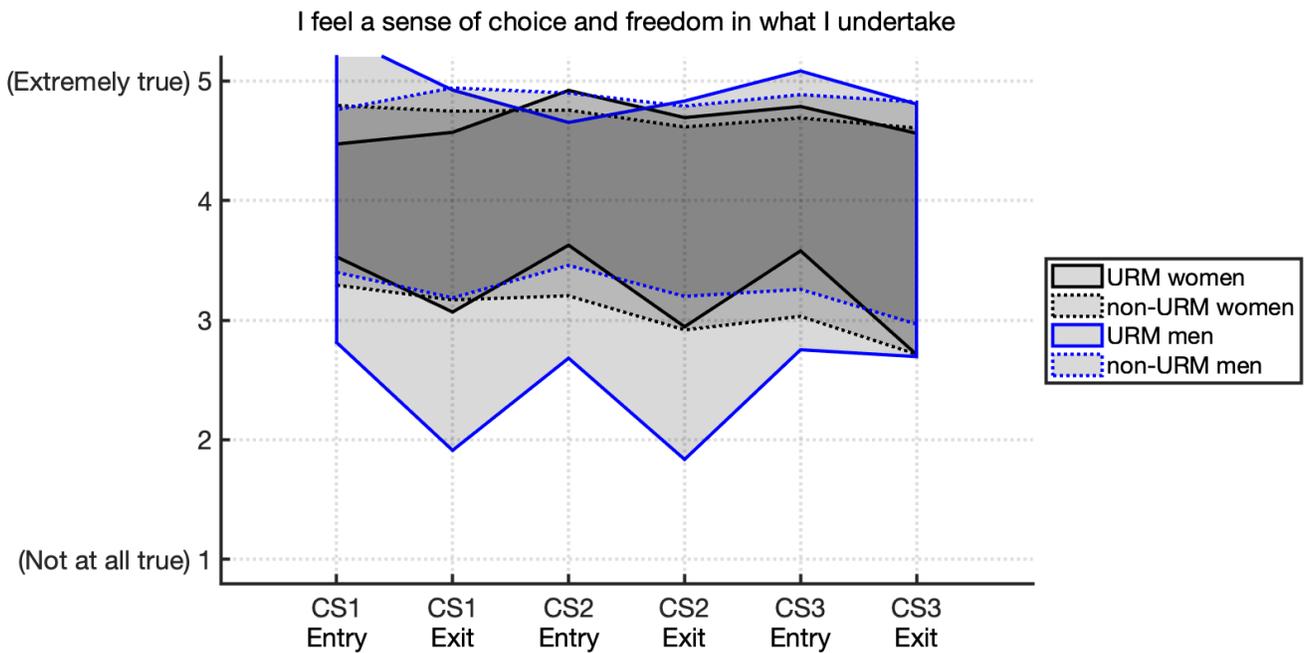


Figure 4: Spread of the responses to Question 1; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

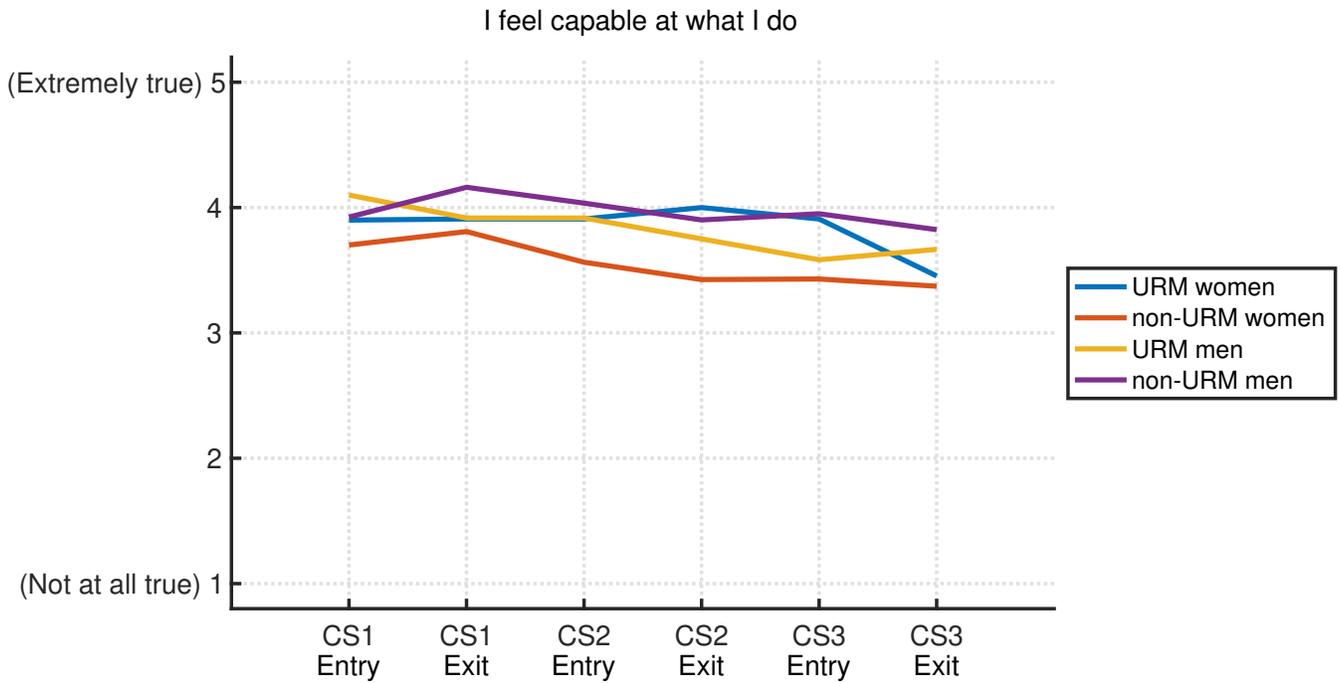


Figure 5: Means of the responses to Question 2.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

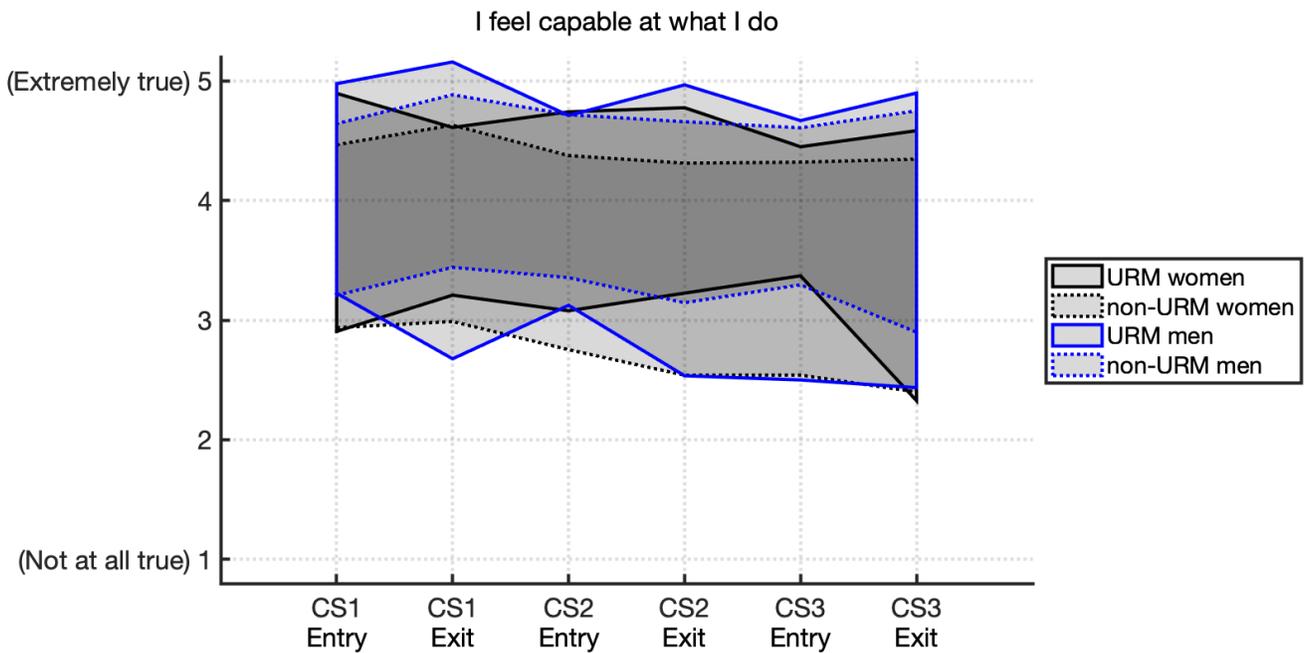


Figure 6: Spread of the responses to Question 2; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

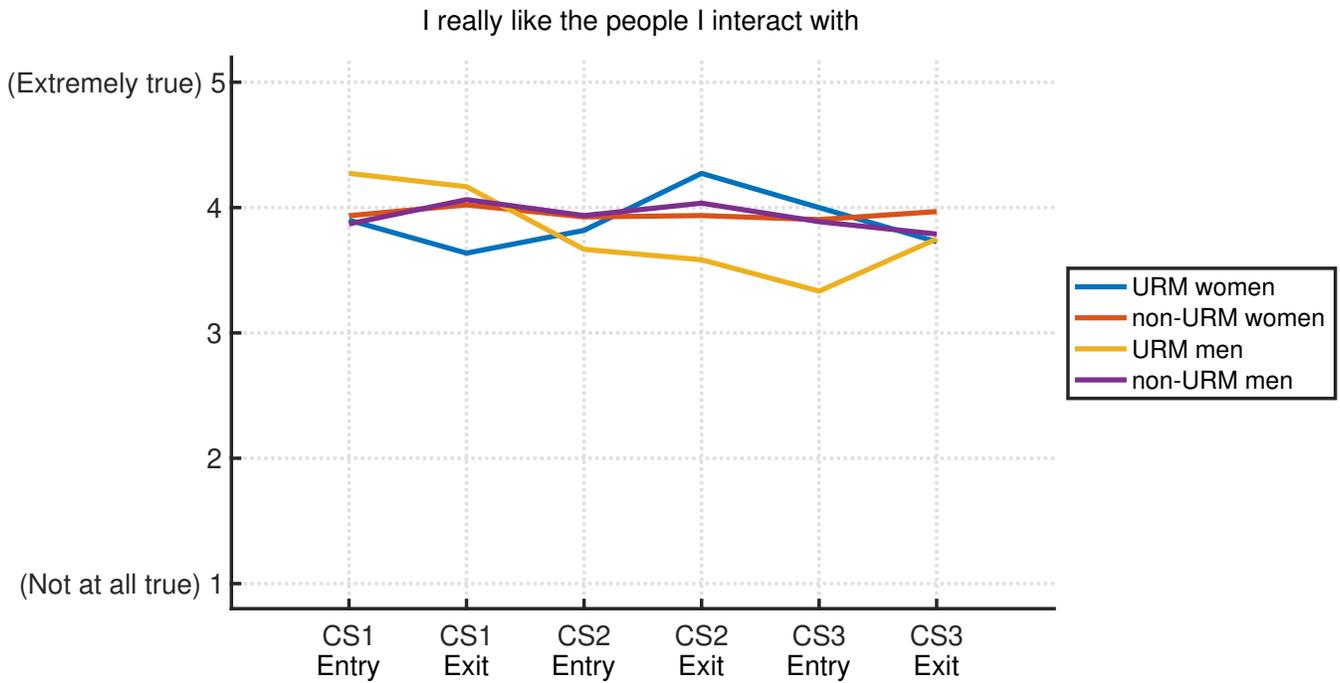


Figure 7: Means of the responses to Question 3.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

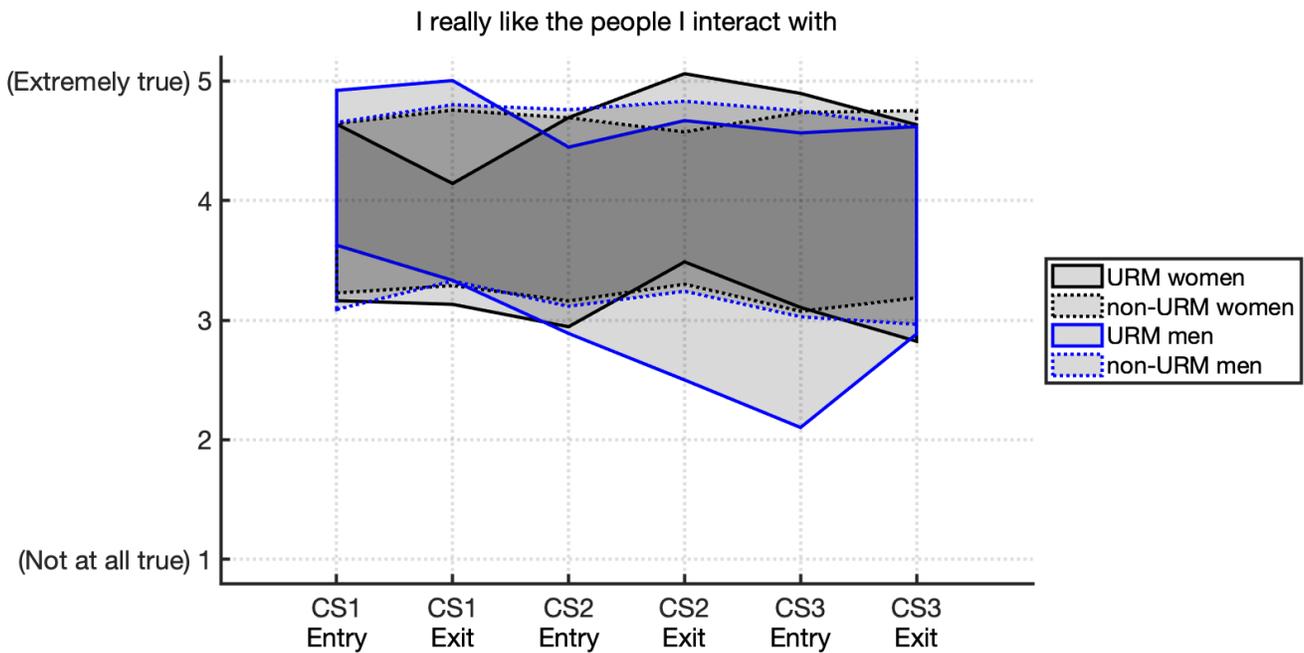


Figure 8: Spread of the responses to Question 3; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

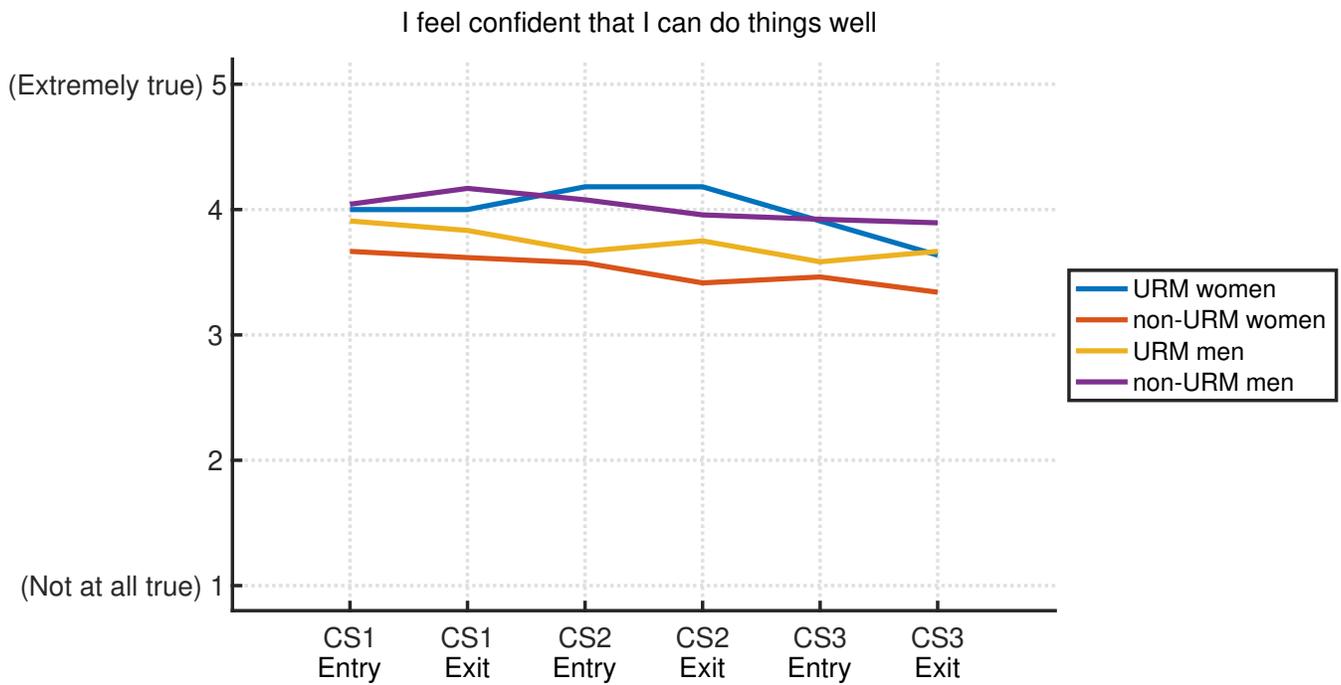


Figure 9: Means of the responses to Question 4.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

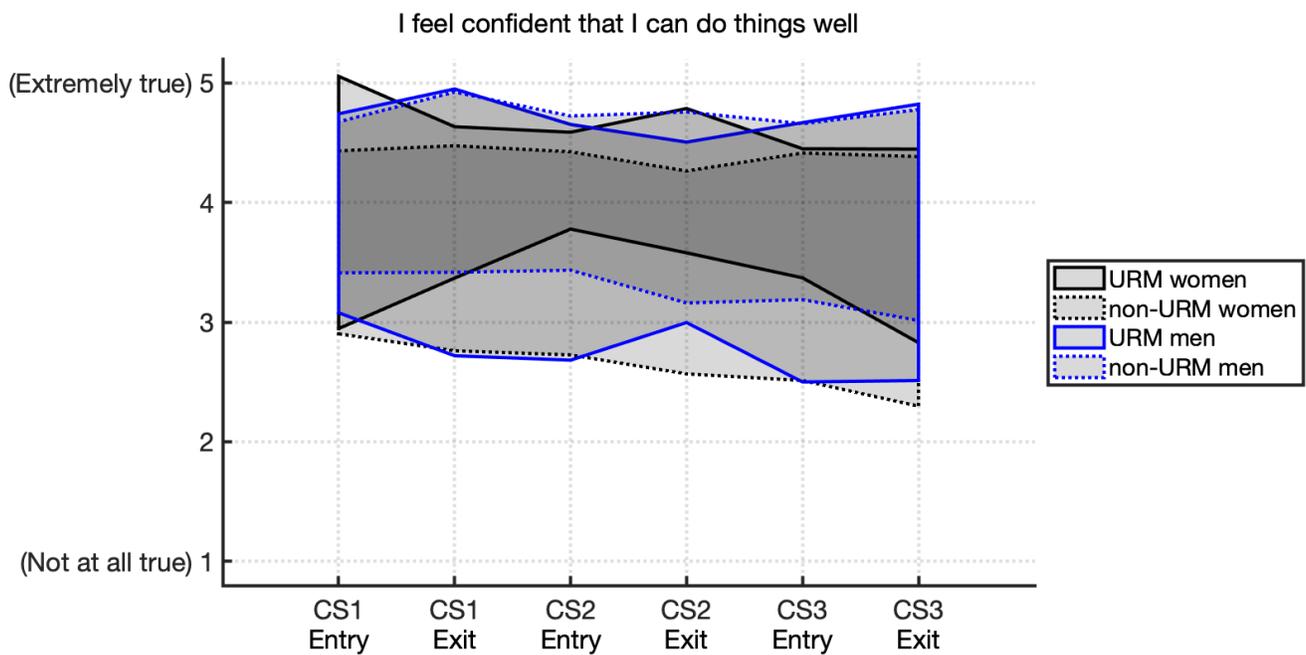


Figure 10: Spread of the responses to Question 4; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

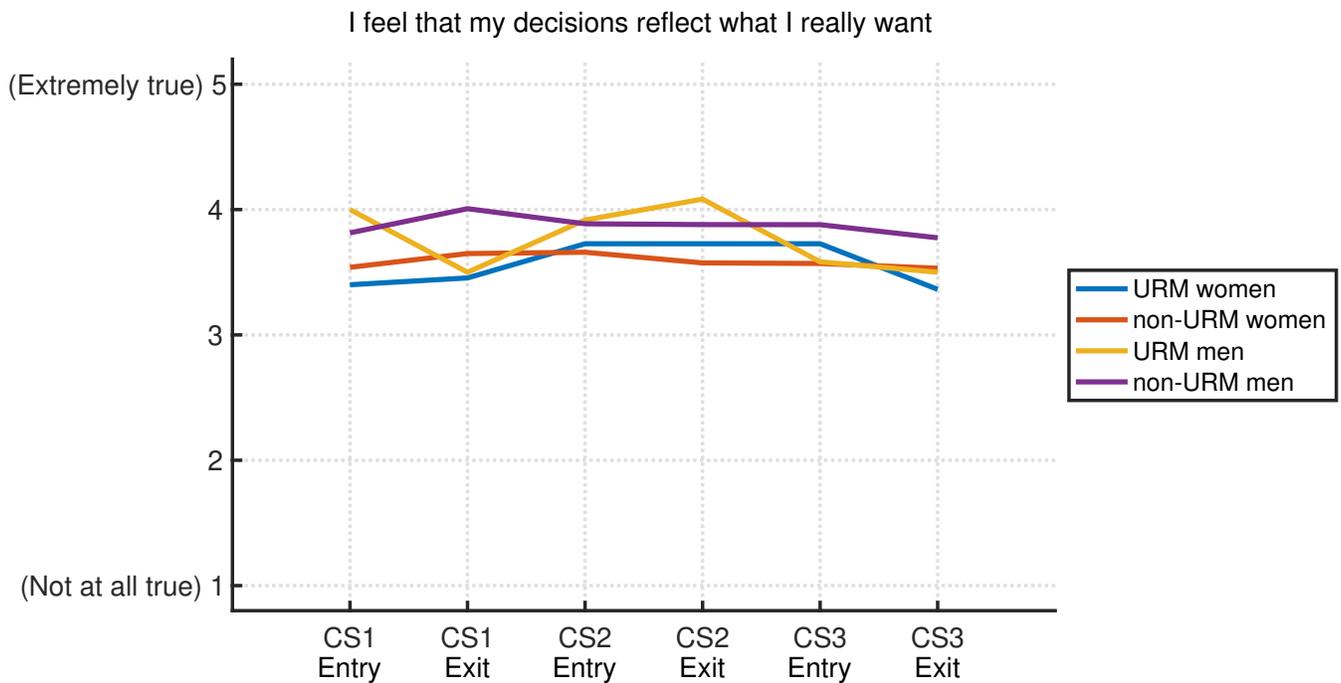


Figure 11: Means of the responses to Question 5.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

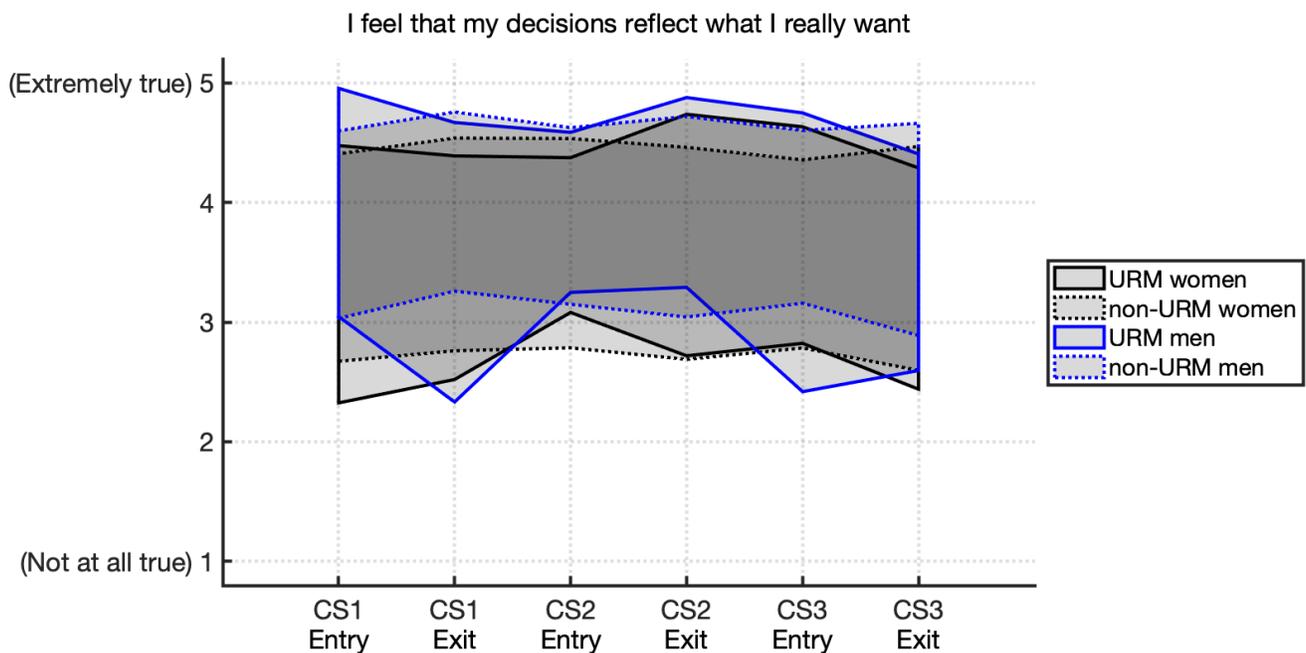


Figure 12: Spread of the responses to Question 5; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

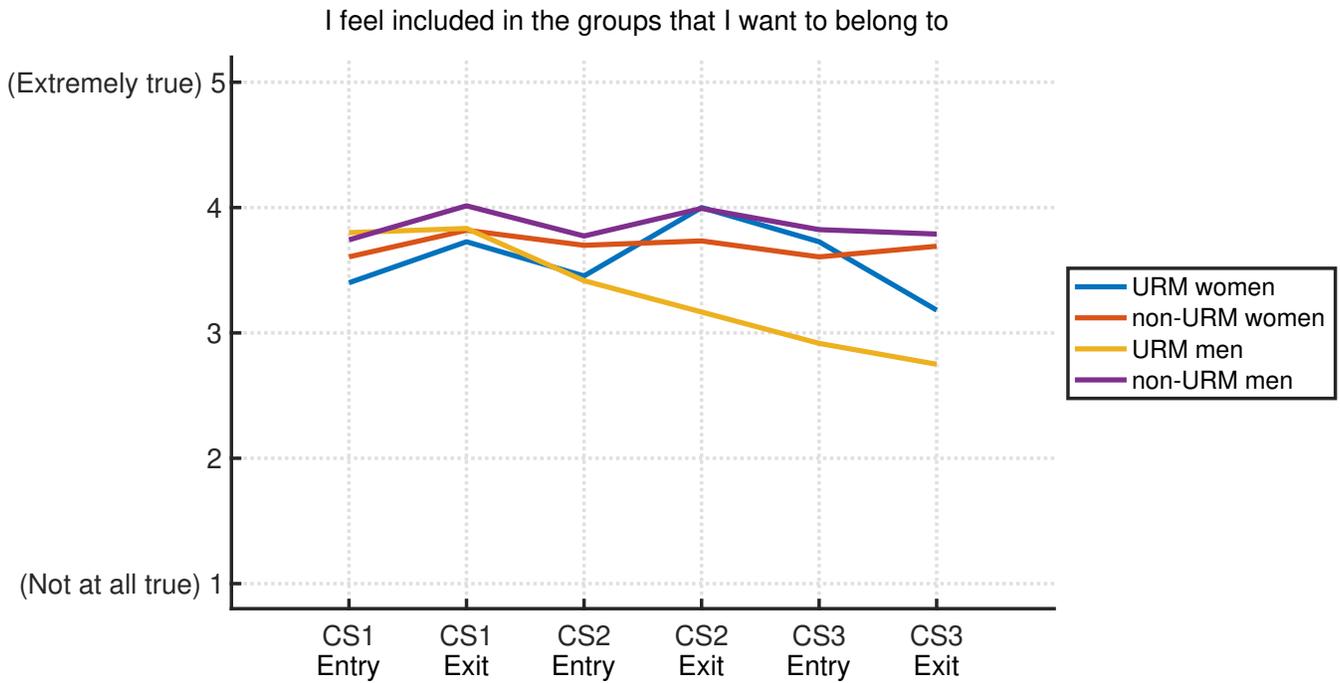


Figure 13: Means of the responses to Question 6.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

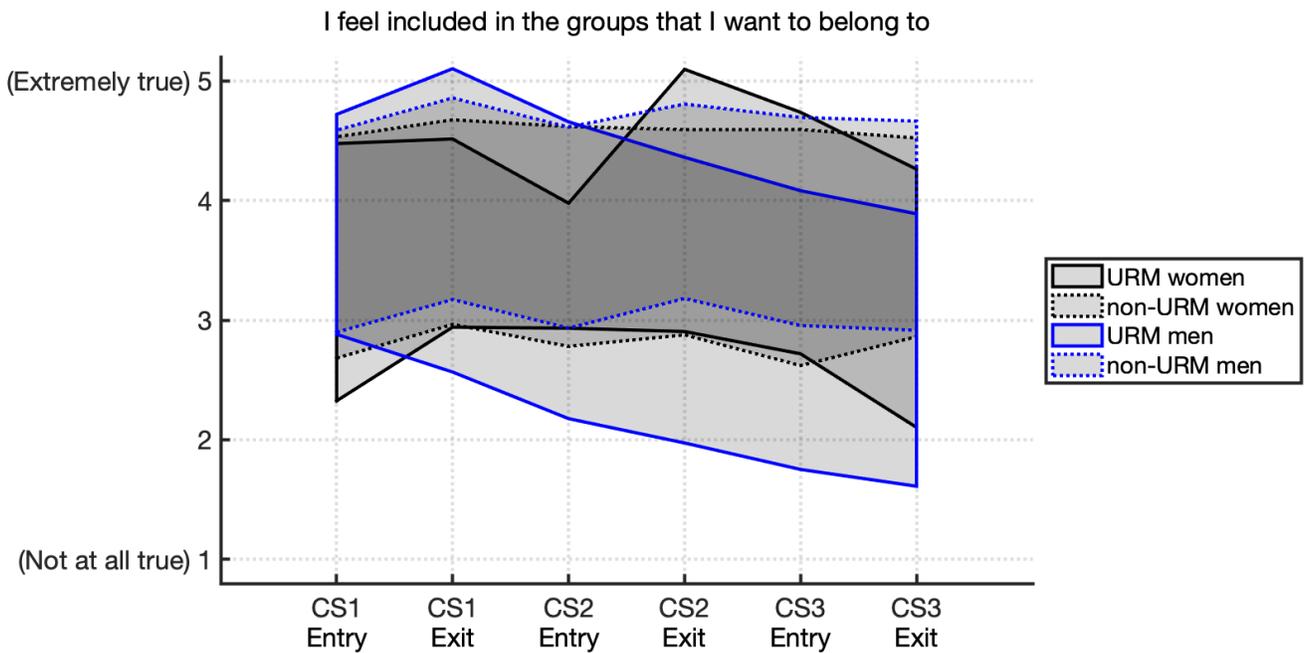


Figure 14: Spread of the responses to Question 6; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

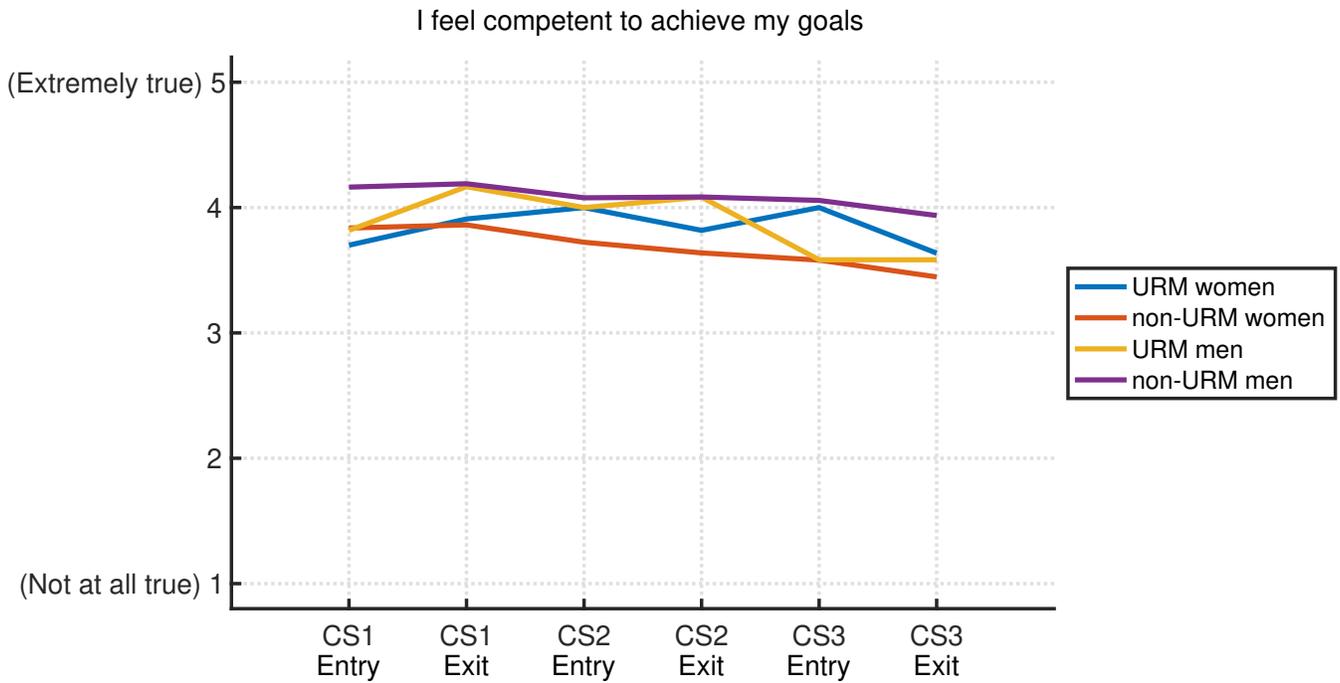


Figure 15: Means of the responses to Question 7.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

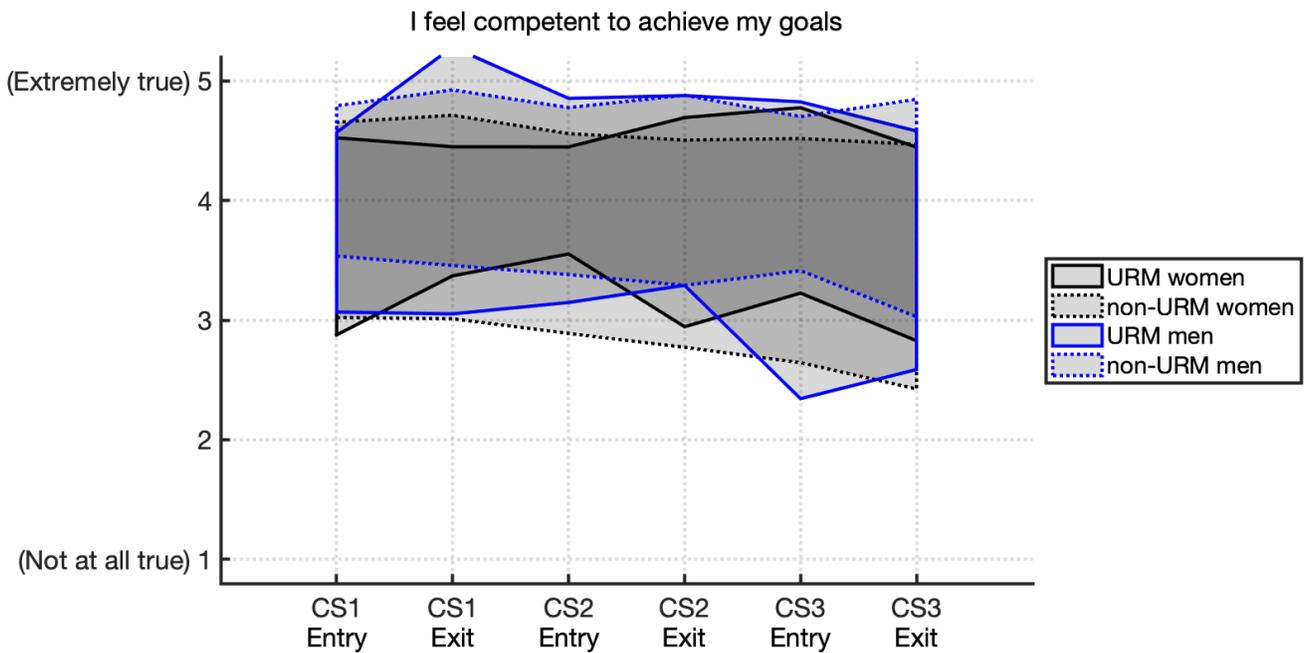


Figure 16: Spread of the responses to Question 7; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

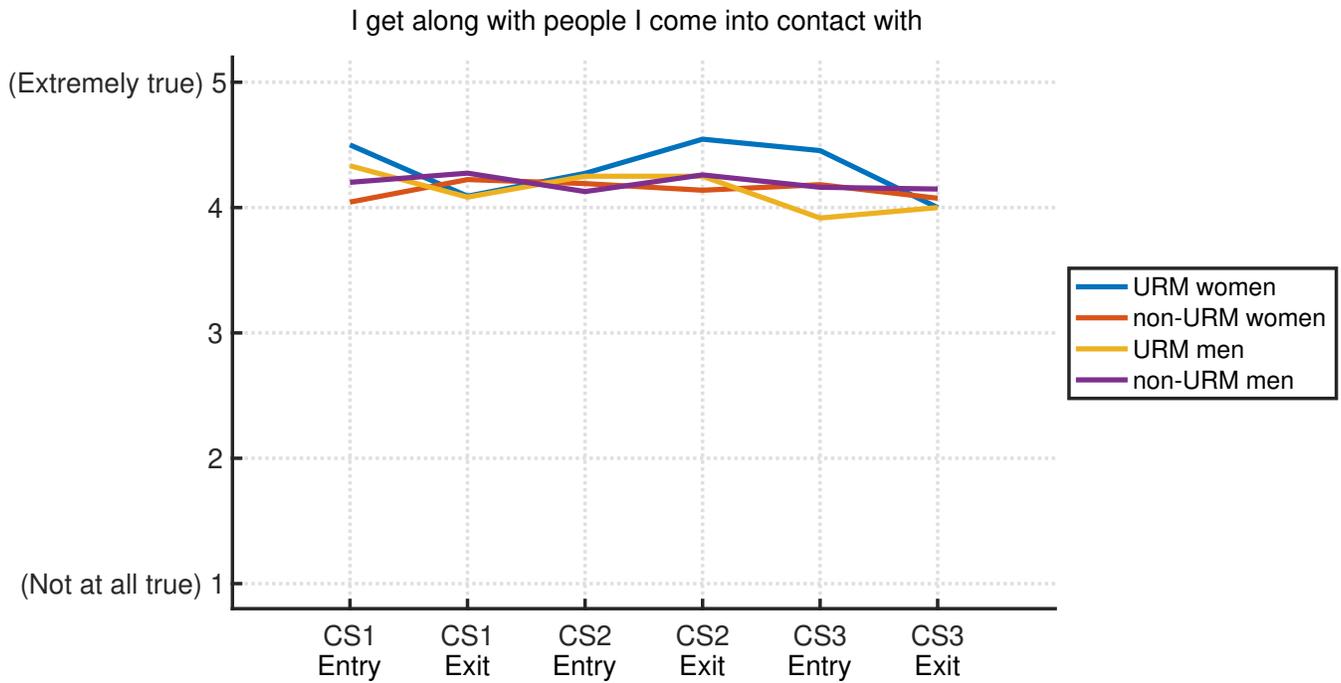


Figure 17: Means of the responses to Question 8.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

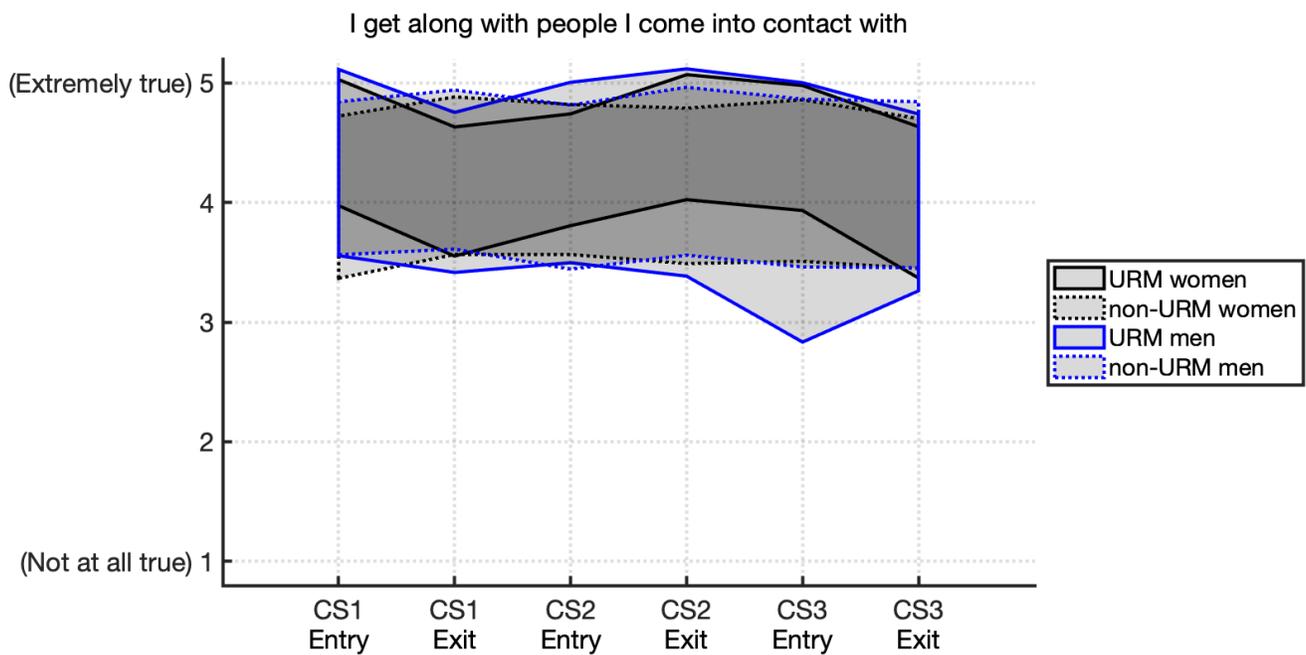


Figure 18: Spread of the responses to Question 8; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

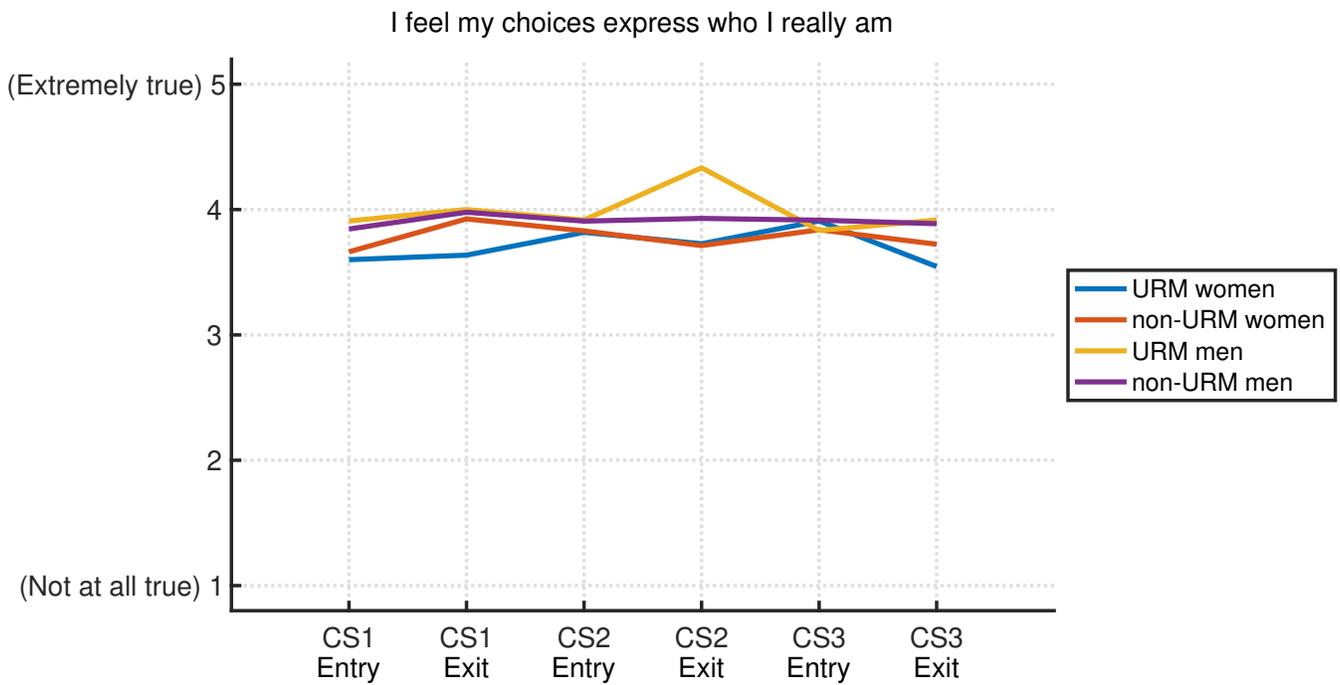


Figure 19: Means of the responses to Question 9.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

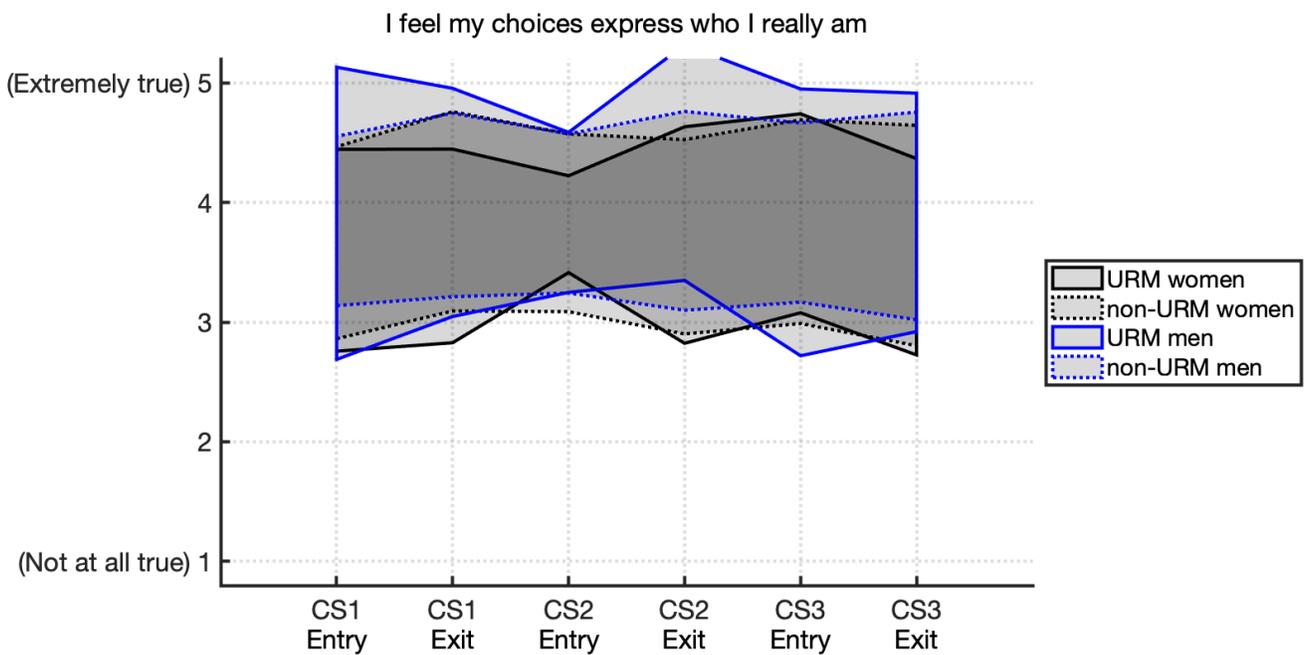


Figure 20: Spread of the responses to Question 9; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

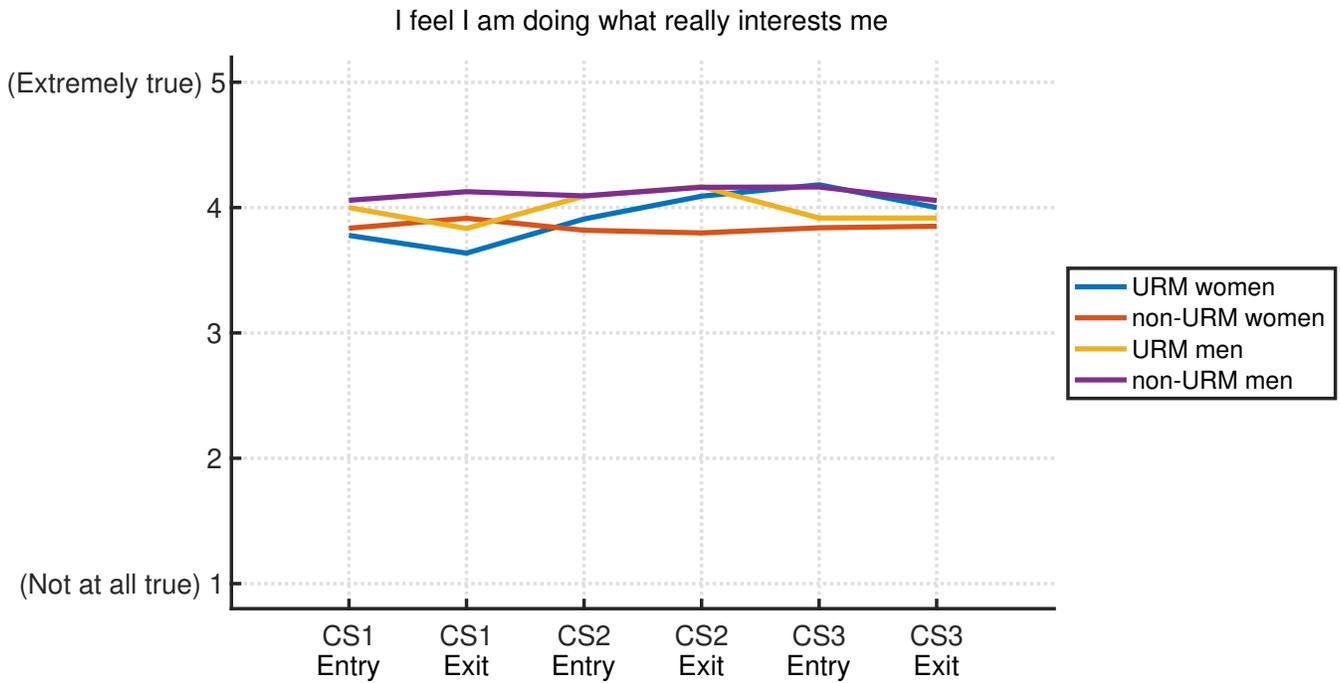


Figure 21: Means of the responses to Question 10.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

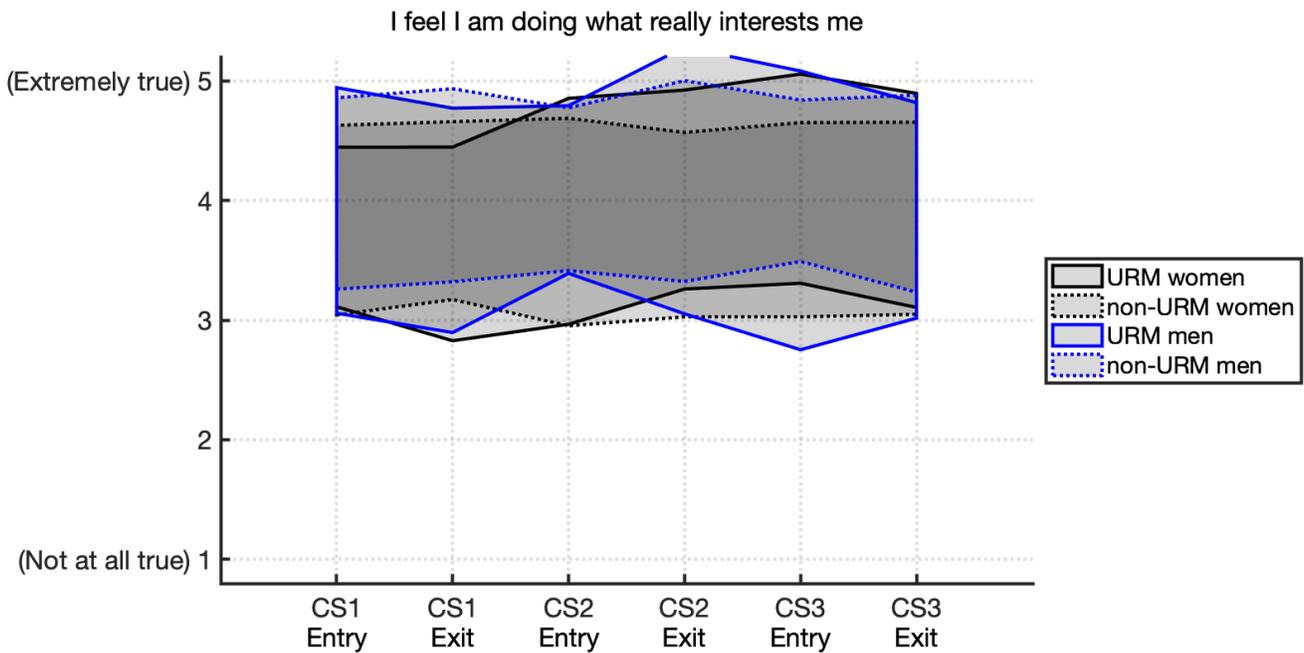


Figure 22: Spread of the responses to Question 10; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

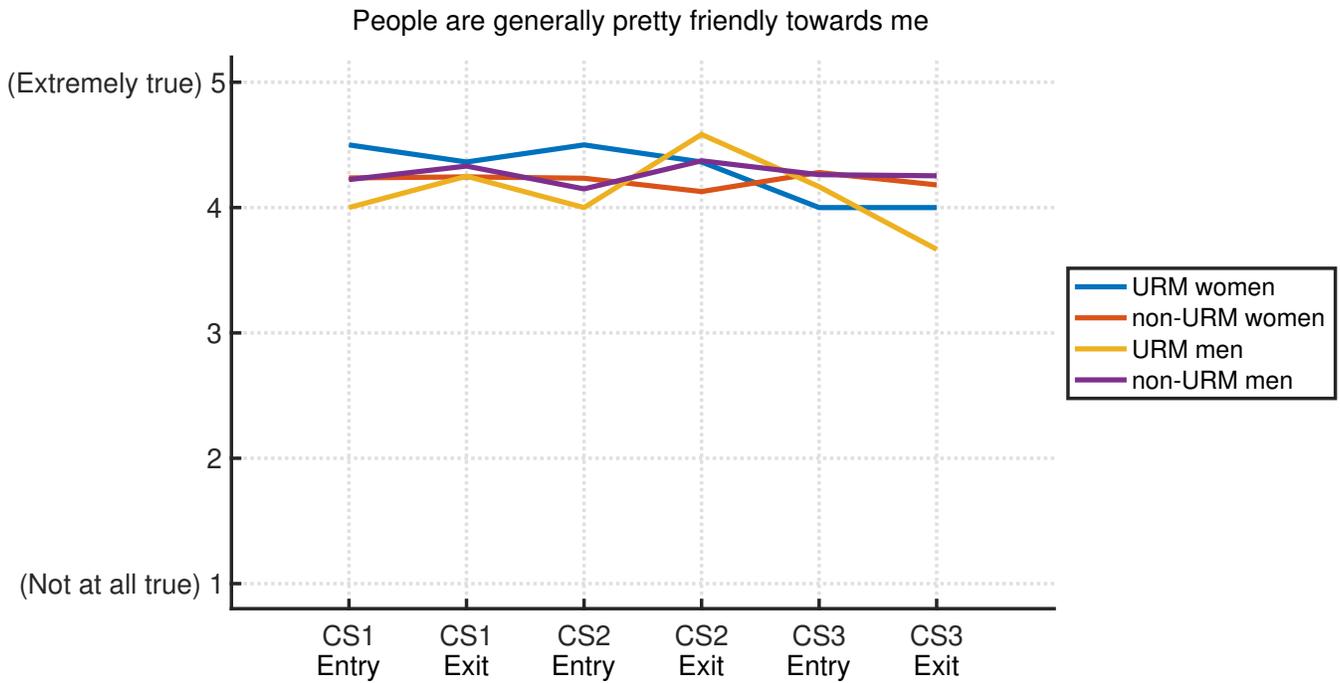


Figure 23: Means of the responses to Question 11.
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

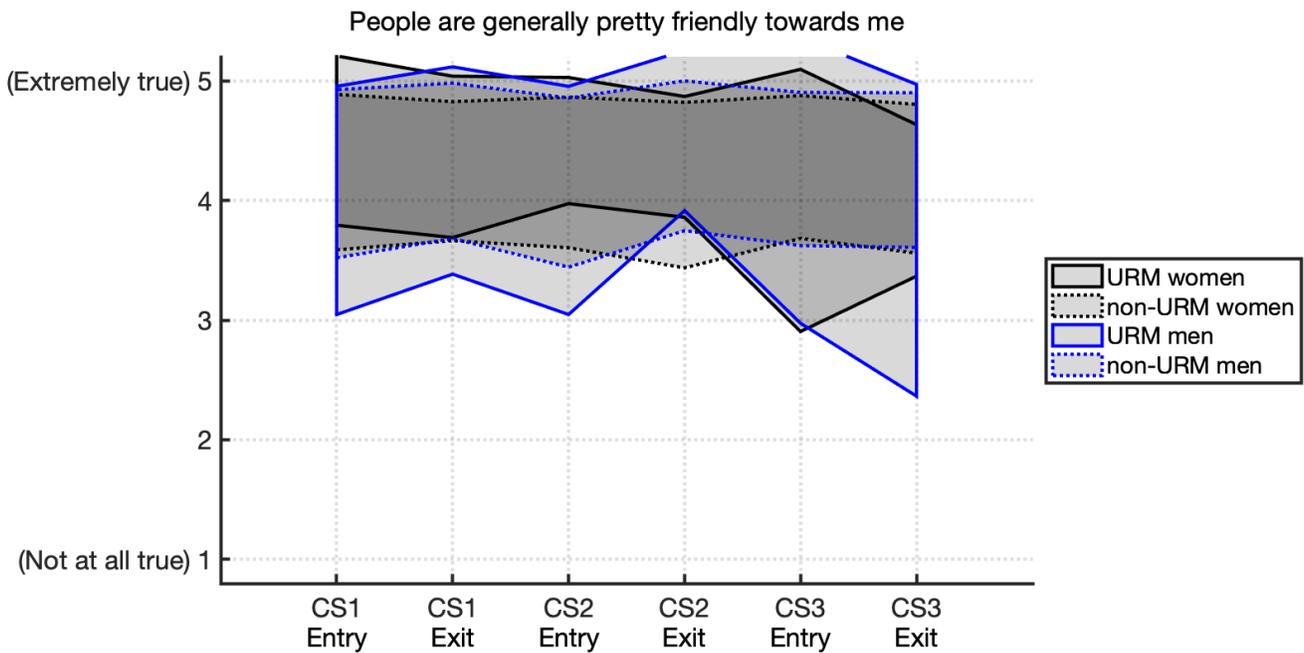


Figure 24: Spread of the responses to Question 11; filled bands are the mean +/- one standard deviation.
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

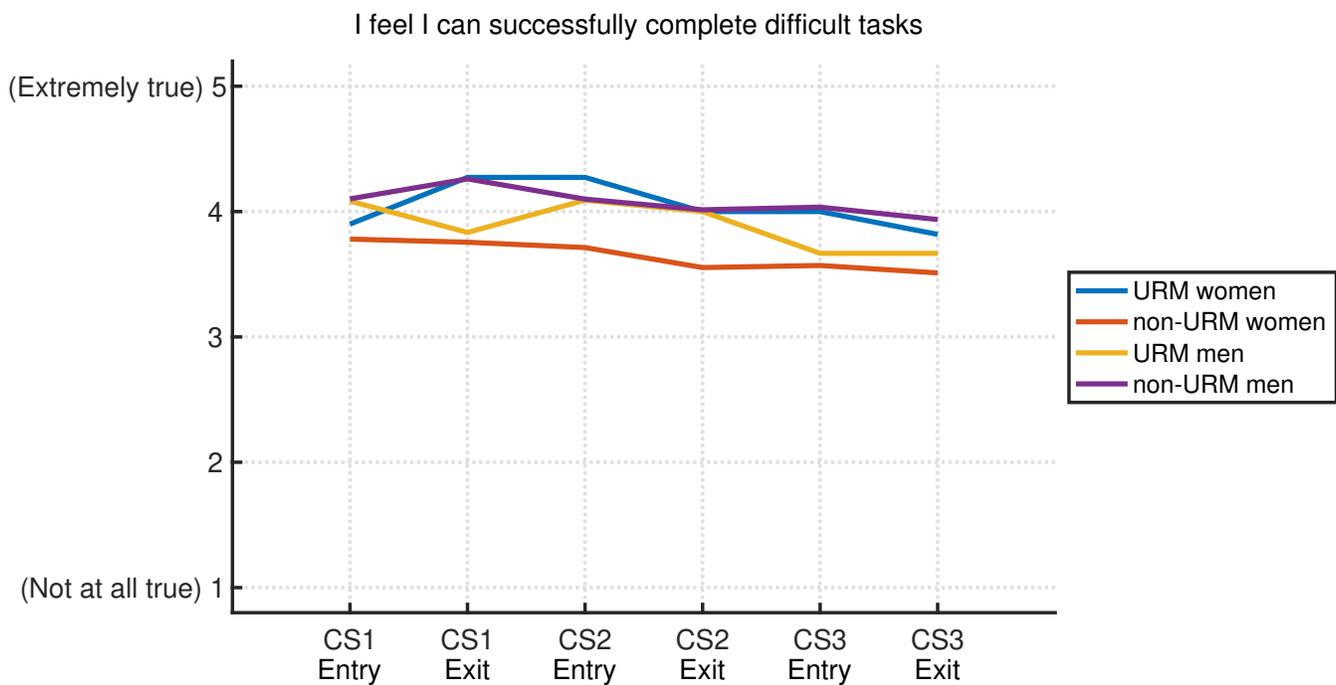


Figure 25: Means of the responses to Question 12.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

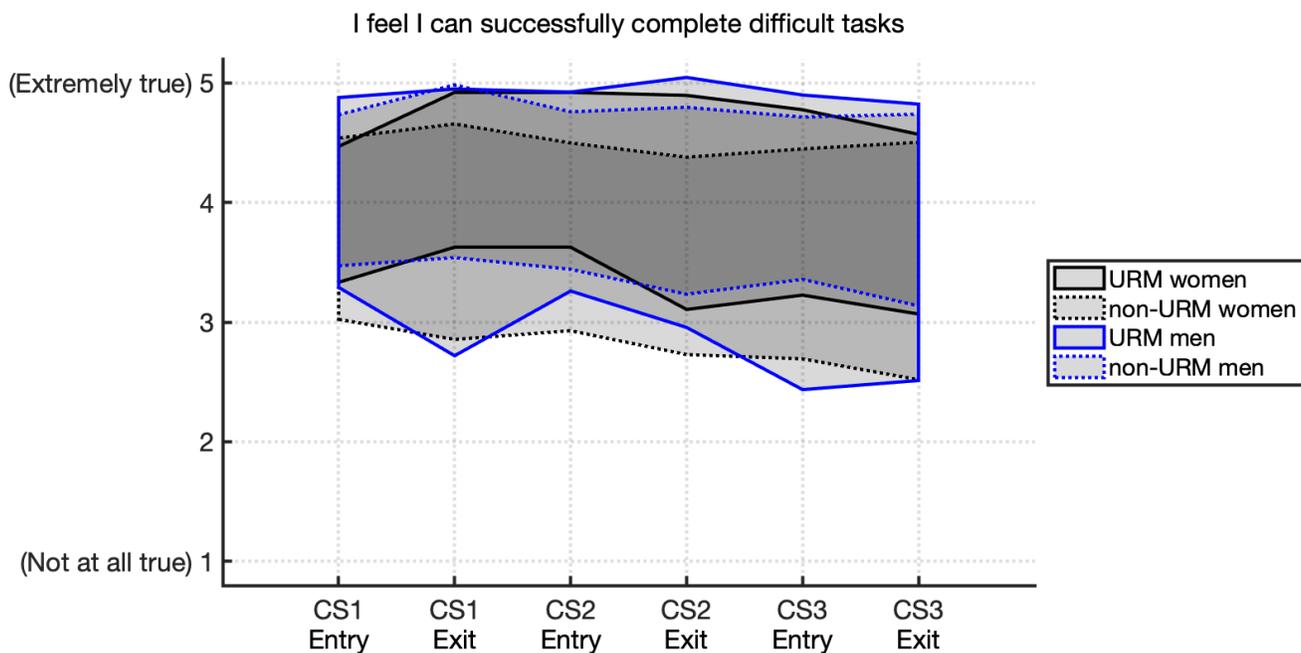


Figure 26: Spread of the responses to Question 12; filled bands are the mean +/- one standard deviation.
N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

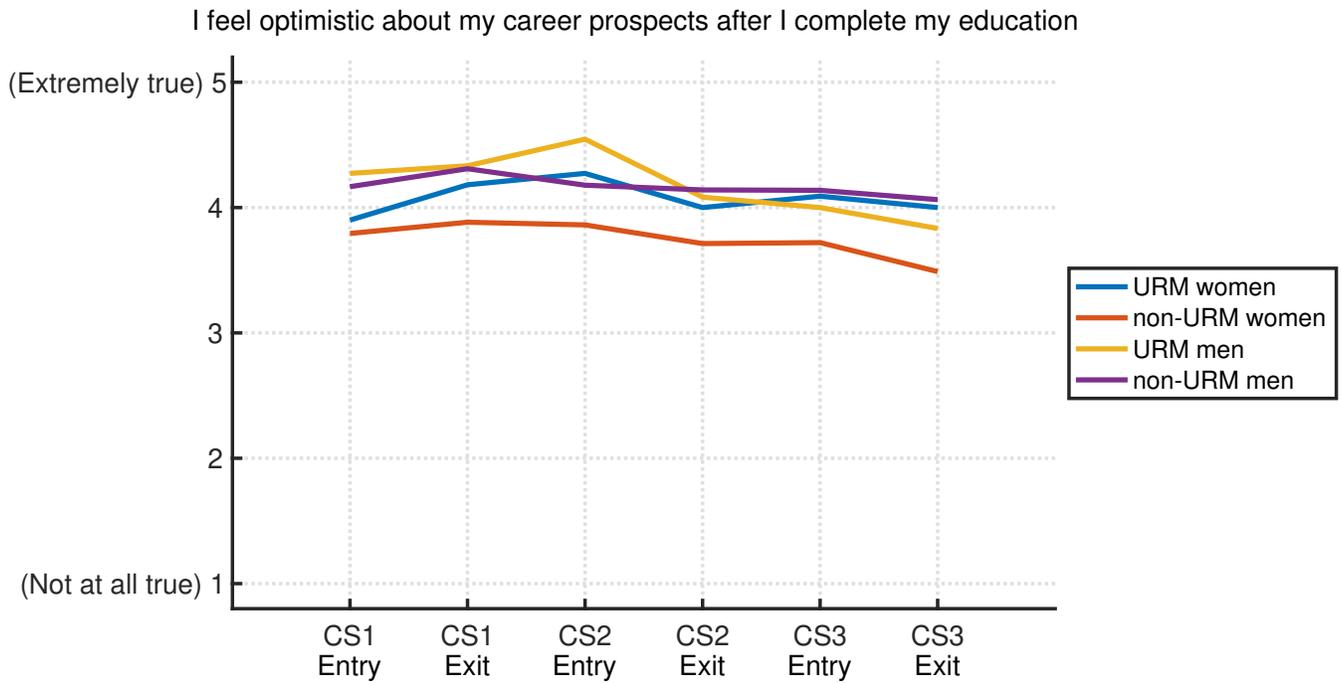


Figure 27: Means of the responses to Question 13.
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.

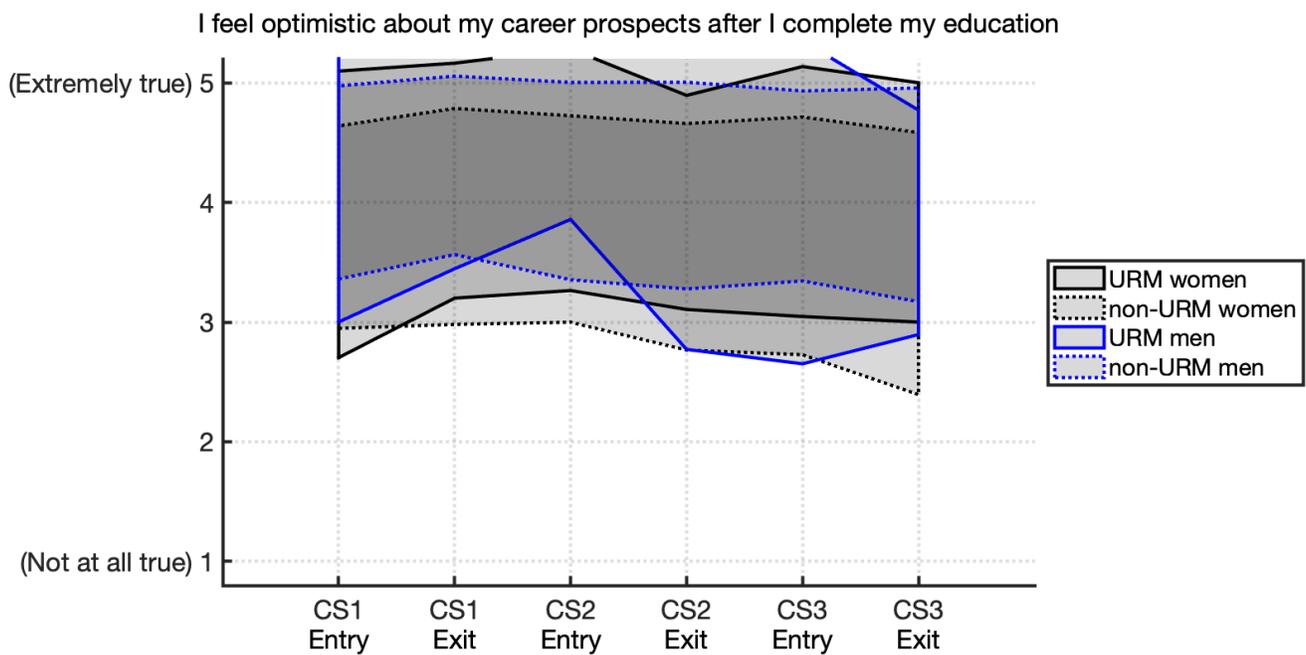


Figure 28: Spread of the responses to Question 13; filled bands are the mean +/- one standard deviation.
 N = 259; 11 URM women, 94 non-URM women, 12 URM men, 142 non-URM men.