Student Perceptions of Their Abilities and Learning Environment in Large Introductory Computer Programming Courses - Underrepresented Minorities

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. Andrew DeOrio, University of Michigan

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 Perceptions of Their Abilities and Learning Environment in Large Introductory Computer Programming Courses – Under-Represented Minorities

Abstract

Historically, institutions have struggled to increase the number of underrepresented minority (URM) students completing computer science and computer engineering undergraduate degrees.

There are many potential obstacles to student success. Faculty that teach the large programming courses at our university identified three particular obstacles to diversity in computer science and computer engineering: stereotyped traits, perceived abilities, and learning environment. Identifying implicit bias and imposter syndrome as components of these obstacles, we implemented a series of 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. In particular, we are interested in the difference between URM students' and non-URM students' perceptions of their abilities and the learning environments in these courses.

In the present study, our overarching research question is: Do underrepresented students and non-underrepresented students 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 presents entry and exit survey results from three semesters (Fall 2017, Winter 2018, and Fall 2018) of two successive programming courses. The results were analyzed using mixed model ANOVA for repeated measures of questions on self-efficacy, intimidation by programming, and feelings of inclusion.

Statistically significant results include: We observed a decrease in self-efficacy during the term for both courses in our study, although the effect is small and the decrease is slightly larger for URM students than non-URM students in Course 1; and a decrease in inclusion for students in Course 1, though again the effect is small. Overall, the perceptions of URM students are similar to non-URM students.

Introduction and Motivation

A recent survey by the National Science Foundation showed that higher education is still struggling to increase the number of underrepresented minority (URM) students completing computer science and computer engineering undergraduate degrees[1]. The reasons for this URM gap are complex, but they can be generalized into the problems of recruitment and retention.

Recruiting URM students to computer computer science and computer engineering can be difficult because high school students have preconceptions ideas about who does – and does not – go into computer science and computer engineering.[2] Retaining URM students can be challenging because it involves the behaviors and attitudes of many different instructors and hundreds or thousands of classmates.[3,4]

There are many potential obstacles to recruiting and retaining URM students, and ultimately these obstacles are also obstacles to individual student success in computer science and computer engineering. Faculty that teach the large programming courses at our university identified three particular obstacles to diversity in computer science and computer engineering: stereotyped traits,[5] perceived abilities,[6] and learning environment.[7,8] Identifying implicit bias and imposter syndrome as components of these obstacles, we implemented a series of activities designed to lessen the impact of implicit bias and imposter syndrome on our students in large-enrollment introductory computer programming courses.

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, imposter syndrome, and stereotype threat
- student activities related to implicit bias and imposter syndrome
- personalized messaging via an electronic coaching system

Our previous work[9] looked at ways to remove or minimize the impact of three obstacles to diversity in the computer science and computer engineering undergraduate programs: stereotyped traits, perceived abilities, and learning environment. Originally, the study was focused on obstacles to gender diversity, but these obstacles 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. Therefore, we embarked on a 5 year program to gather data and assess the differences in student perceptions in the large programming courses offered at our institution.

Our initial analysis of gender differences in student perceptions of the impact of stereotypes, preconceived notions of ability, and learning environment on their experiences in their programming courses showed that there was a statistically significant difference between men and women's perceptions, though the overall differences were small.[10] We turn now to comparing the responses of underrepresented minority (URM) students vs. non-URM students.

Recruiting and retaining 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 URM and non-URM students 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 courses. Our hope was that URM students' perceptions of their abilities and learning environment would show a positive change from the entry survey to the exit survey if our implicit bias and stereotype threat activities were effective. 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 two different large-enrollment computer programming courses at a large, public research institution. 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 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 who opted in to the research program 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 for the Fall 2018 term is show in Fig. 1, and understanding the material, learning, and grades were all considered aspects of success.

Figure 1: Word cloud resulting from student responses to the question, "Briefly describe what being SUCCESSFUL in [Course 1] means to you." Understanding the material, learning, and grades were all considered metrics by which students measure their own success.

Self-efficacy can affect the grade a student earns;[11] 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.[12] 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 URM students' self-efficacy over the course of the term.

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

Indicator 2: Intimidation. Intimidation is the summation of many things. Women on average are at a disadvantage when it comes to when they were first introduced to computers;[13] what experiences, role models, and encouragement they had in K-12,[14] and previous programming experience.[15,16,17] These experiences, taken together, can lead to an increased feeling of intimidation when women are required to take a programming course. Based on our previous analysis of men's vs. women's perceptions of intimidation in programming courses, we hypothesized that URM students have a similar level of incoming intimidation, even though they may have more previous programming experience than women, on average.[18] As we do not have any control over our students' previous experiences about programming, 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: URM students are more intimidated by programming in our programming courses, as compared to their non-URM 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 URM students will feel welcome and valued in a field[19] and hence choose to major in that field. Peer encouragement[20,21] and respect[22] 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 a increase in students' perceptions of being welcome over the course of the term, with URM students reporting a larger increase over the course of the term.

> Hypothesis 3: URM students feel less welcome in our programming courses, as compared to their non-URM 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 16.

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

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

Analysis

We have two within-subjects repeated measures: the start of term and the end of term. There are two between-subjects measures: URM students and non-URM students. 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 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 that identify as African-American/Black, Hispanic, and Native American. In

our analysis, we categorized all students who identified as solely "White" or solely "Asian" as non-URM students and all other students as URM students.

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 non-URM students' responses on the left and URM students' responses on the right. The y axis is labeled with the responses (linear scale or multiple choice); the x axis is the percentage of non-URM or URM students 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 16.

Table 2: Summary of responses for those students who took *both* the entry and the exit surveys in a given term. Three terms' worth of data are combined here: Fall 2017, Winter 2018, Fall 2018.

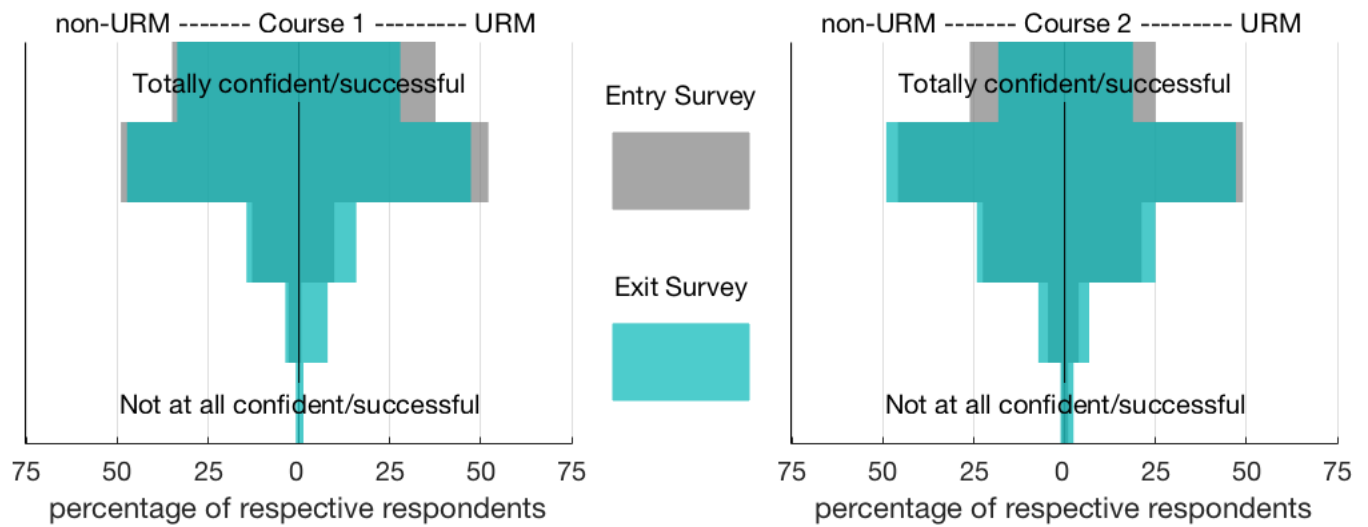| Course	Enrollment	Total # Respondents	Response Percentage	URM Respondents		Non-URM Respondents	
				#	%	#	%
1	2085	932	44.7%	165	17.7%	767	82.3%
2	2907	1766	60.8%	208	11.8%	1558	88.2%
Total	4992	2698	54.05%	373	13.8%	2325	86.2%

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)
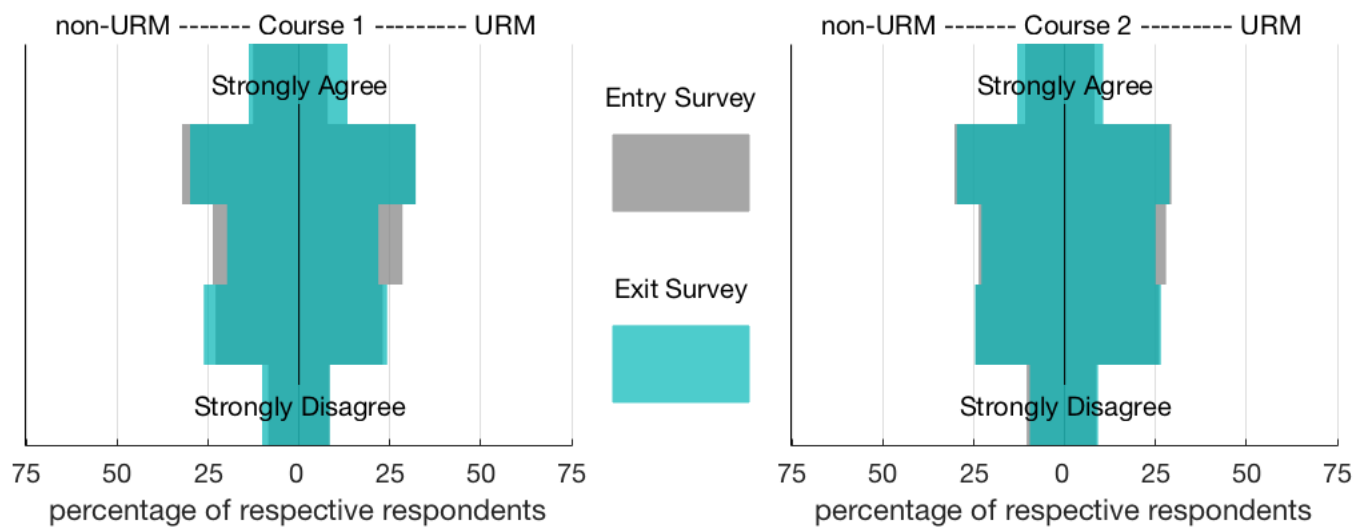

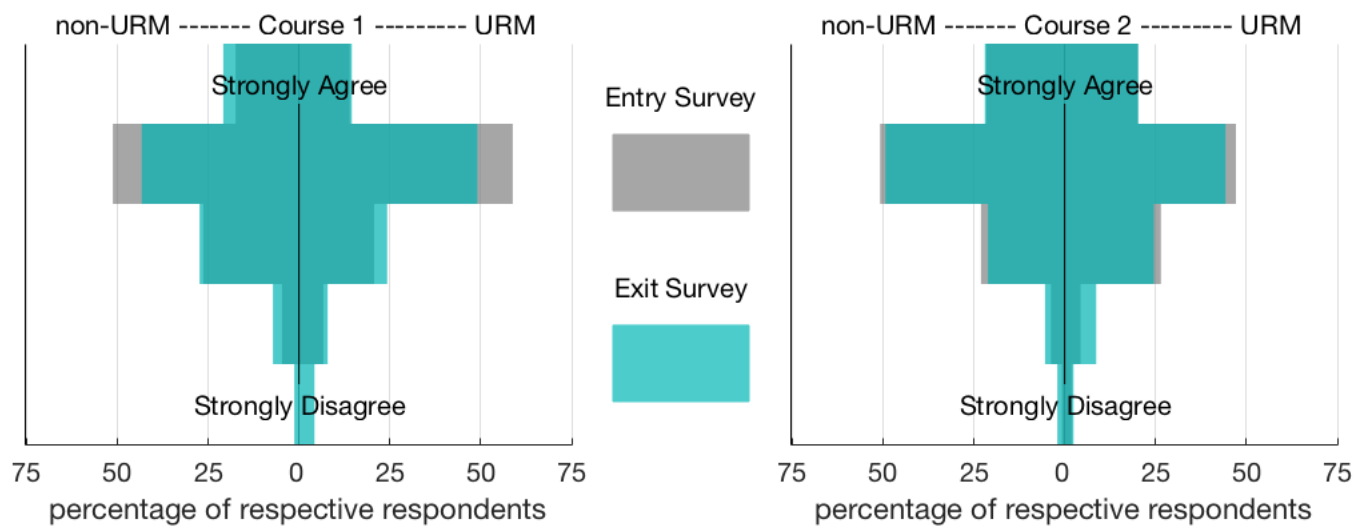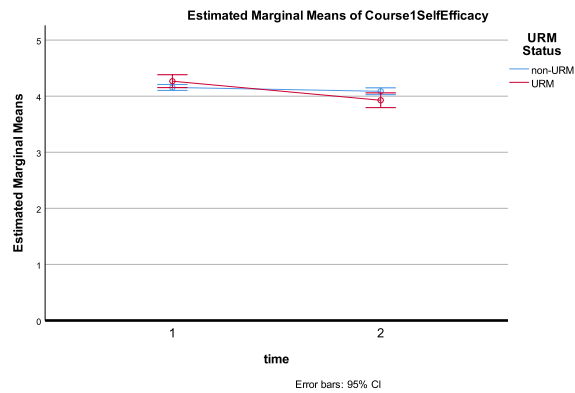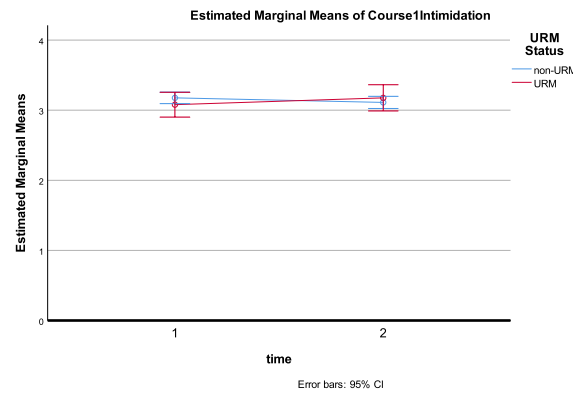Figure 3: Distribution of responses to the statement, "I find computer programming intimidating."

Figure 4: Distribution of responses to the statement, "I believe that other students in computer programming courses will be welcoming of me."

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	4.16	0.770	767
	URM	4.27	0.655	165
	Total	4.17	0.752	932
exitSurvey	non-URM	4.09	0.841	767
	URM	3.93	0.928	165
	Total	4.06	0.858	932

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.18	1.169	767
	URM	3.08	1.099	165
	Total	3.16	1.157	932
exitSurvey	non-URM	3.11	1.229	767
	URM	3.18	1.189	165
	Total	3.12	1.222	932

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.81	0.783	767
	URM	3.80	0.759	165
	Total	3.81	0.779	932
exitSurvey	non-URM	3.75	0.909	767
	URM	3.62	0.972	165
	Total	3.73	0.921	932

(a) Self-Efficacy. Within-subjects (time) is statistically significant, as is time * URM Status.

(b) Intimidation by Programming. No statistical significance.

(c) Inclusion. Within-subjects (time) is statistically significant.

Figure 5: Summary of mixed model ANOVA results for Course 1. See Appendix for more details.

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.92	0.860	1558
	URM	3.93	0.837	208
	Total	3.92	0.857	1766
exitSurvey	non-URM	3.76	0.873	1558
	URM	3.73	0.925	208
	Total	3.76	0.879	1766

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.07	1.184	1558
	URM	3.02	1.110	208
	Total	3.06	1.175	1766
exitSurvey	non-URM	3.12	1.200	1558
	URM	3.05	1.160	208
	Total	3.11	1.195	1766

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.89	0.809	1558
	URM	3.79	0.879	208
	Total	3.88	0.818	1766
exitSurvey	non-URM	3.84	0.899	1558
	URM	3.71	0.965	208
	Total	3.82	0.908	1766

(a) Self-Efficacy. Within-subjects (time) is statistically significant.

(b) Intimidation by Programming. No statistical significance.

(c) Inclusion. No statistical significance.

Figure 6: Summary of mixed model ANOVA results for Course 2. 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 16 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 URM and non-URM students report similar feelings of self-efficacy in both courses (the histogram view in Fig. 2 shows near symmetry left/right) and that in general their feelings on self-efficacy are high (~ 4 on a 5 point scale), although there is a small decrease between the beginning of the term and the end of the term. There is no statistical significance between subjects (URM vs. non-URM). Within-subjects (time) is statistically significant; this means that the decrease in self-efficacy over the course of the term happens across the whole sample, although the size of the effect is small. Within-subjects (time * URM Status) is also statistically significant for Course 1, though not Course 2; this means that the decrease in self-efficacy is not equivalent across the two groups (URM and non-URM students) in Course 1. As we see in 5a, URM students do report a larger decrease in self-efficacy, but again the size of the effect is small. Although there is statistical significance for time * URM Status, we cannot reject the null hypothesis for Hypothesis 1 because we hypothesized that feelings of self-efficacy increased, and instead they decreased slightly.

Intimidation
Visually inspecting the survey results for intimidation by programming in Fig. 3, we see that overall both URM and non-URM students are somewhat evenly spread in how intimidated they feel by programming and URM and non-URM students are similar in their reporting on intimidation (Fig. 3 is fairly symmetrical overall, and between-subjects is not statistically significant). There are small shifts in intimidation between the entry and exit surveys, as shown in Figs. 3, 5b, and 6b, but these shifts (within-subjects: time and time * URM Status) are not statistically significant. We cannot reject the null hypothesis for Hypothesis 2 because there was not a statistically significant difference between URM and non-URM students nor was there a decrease in reported intimidation.

Inclusion
Visually inspecting the survey results for inclusion in Figs. 4, we see a shift from entry survey to exit survey for Course 1, and less of a shift for Course 2.

The marginal means for URM and non-URM students in Course 1 are almost identical for the entry survey, but the distribution is curiously different. More URM students report "agree" for inclusion, which is rated as a 4 and is close to the mean value of 3.80. The non-URM students, which have a mean value of 3.81, have less responses at "agree" and more at both "strongly agree" and "neutral" – resulting in a similar mean to the URM students but indicating a different perspective from the non-URM group. It is possible that this is due to the number of women's responses being in the neutral category, as reported on in a previous study.[10]

Non-URM students in Course 1 show a shift of students feeling both more welcome (more responses of "strongly agree" on the exit survey) and less welcome (more responses of "disagree"

and "strongly disagree" on the exit survey). This mirrors the response of men overall that was found in the previous study cited above (women showed little change on inclusion for Course 1). URM students in both courses show a small decrease in feelings of inclusion.

Looking more closely at the data in Figs. 5c, and 6c, the within-subjects (time) was statistically significant only for Course 1. Between-subjects (URM and non-URM) was not significant. Therefore, we cannot reject the null hypothesis for Hypothesis 3.

Overall Interpretation
These results indicate that URM and non-URM students in these courses have very similar perceptions of self-efficacy, intimidation, and inclusion when considering computer science. And in general, these perceptions are favorable: fairly high self-efficacy, moderate amounts of intimidation, and fairly good inclusion.

In terms of the raw numbers, we are fairly satisfied with the reported perceptions of self-efficacy. While we would like to lower the feelings of intimidation, the spread that we see in the histograms in Fig. 3 may be the best we can do until computer programming becomes a more universally taught skill in primary and secondary schools. Inclusion, in particular, is not as high as we would like to see for the entire group of students. This is certainly an area that we can work to improve.

Overall, these results seem promising for our courses, but our findings here are contrary to the prevailing understanding that URM students have lower self-efficacy, feel more intimidated by programming, and feel less welcome in their CS courses. We have several thoughts on why this may be so.

We are hopeful that the interventions we do in these classes (balanced teaching staff, activities the students do on implicit bias and imposter syndrome, etc.) are part of the reason that we are not seeing significant differences between URM and non-URM students as measured on self-efficacy, intimidation, and inclusion. Another possible reason our results differ from the research is that the rigorous admissions selection process at our institution may lead to URM students that are particularly resilient. In general, it appears that students with higher grade point averages perceive less racism and discrimination in their college courses[23]. Lower perceptions of racism and discrimination may then lead to higher graduation rates,[24] creating a nice feed-forward loop for URM students provided they have the support and opportunities needed to earn those high grade point averages. In our case, it may be that the rigorous admissions process self-selects those URM students who already fall into the "higher GPA" category, thus providing us with extraordinary individuals who are are able to overcome these obstacles to learning regardless of the course learning environment.

We also note that these measures are for the groups (URM and non-URM) as a whole; an individual student of any group may still be at-risk of low self-efficacy, high intimidation, and low inclusion depending on their own personal circumstances.

Limitations

A number of factors were beyond our control in this study, and we discuss several threats to validity.

First, our survey was voluntary and the response rate was 54%. It is possible that the respondents were a self-selected group and not representative of the students as a whole.

Second, the students in our study were from a single institution. If the population of students at our competitive, highly ranked public research institution are not representative of students at other institutions, our results could fail to generalize to other universities.

Finally, we do not have a control group. Our program consists of a series of activities designed to lessen the impact of implicit bias and imposter syndrome. This material is presented to all students in our courses. Without a control group, we are not able to compare groups of students exposed to the material against those who are not. However, we do use a within-subjects experiment design to track responses before and after exposure to the material.

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. Students completed pre- and post-surveys in two courses, and our data set includes three semesters.

We presented the statistical analysis for three indicators of student success (self-efficacy, intimidation by programming, and inclusion) in both of the courses in our study. Our first measure was a within-subjects repeated measure: the start of term and the end of term. We also measured the association between-subjects: URM and non-URM students.

We observed that overall, both URM and non-URM students report similar feelings of self-efficacy, intimidation by programming and inclusion. At the end of a course, students generally reported lower measures of all three indicators compared to the beginning. This was a surprising finding, because we expected to see increases as a result of our interventions. We note that the magnitude of these differences were small.

When examining disaggregated data on URM status, we found overall that URM students reported similar levels of self-efficacy, intimidation by programming and inclusion to non-URM students. We noted that on our measure of inclusion, non-URM women often reported "neutral", non-URM men often reported "strongly agree", while URM students mostly responded with "agree". While we did not see an improvement of URM students' perceptions on these three indicators over the course of the term, we are hopeful that our interventions (balanced and diverse teaching staff, discussing implicit bias and stereotype threat, etc.) contribute to minimizing any potential decrease in students' perceptions on self-efficacy, intimidation by programming, and inclusion.

These findings lead to 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 if we are able to capture the whole (or nearly the whole) student population taking these courses? This data only represents roughly half of the students that took these courses.

- 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 [25]), that have a measurable effect on students' perceptions and grades in these programming courses?

Acknowledgement

References

[1] Bachelor's degrees awarded: 2004-2014. Women, minorities, and persons with disabilities in science and engineering. *NSF*, 2017. URL
https://www.nsf.gov/statistics/2017/nsf17310/static/data/tab5-3.pdf.

[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] C. M. Bright, C. A. Duefield, and V. E. Stone. Perceived barriers and biases in the medical education experience by gender and race. *Journal of the National Medical Association*, 90(11):681 – 688, 1998.

[4] L. C. Holley and S. Steiner. Safe space: Student perspectives on classroom environment. *Journal of Social Work Education*, 40(1):49 – 64, 2005.

[5] C. M. Steele. A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6):613–629, 1997. URL http://dx.doi.org/10.1037/0003-066X.52.6.613.

[6] A. J. Gomes, A. N. Santos, and A. J. Mendes. A study on students' behaviours and attitudes towards learning to program. *Proceedings, 17th ACM Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE)*, 2012.

[7] Jennifer M. Grossman and Michelle V. Porche. Perceived gender and racial/ethnic barriers to stem success. *Urban Education*, 49(6):698–727, 2014. doi: 10.1177/0042085913481364. URL
https://doi.org/10.1177/0042085913481364.

[8] Joan Peckham, Lisa L. Harlow, David A. Stuart, Barbara Silver, Helen Mederer, and Peter D. Stephenson. Broadening participation in computing: Issues and challenges. *SIGCSE Bull.*, 39(3):9–13, June 2007. ISSN 0097-8418. doi: 10.1145/1269900.1268790. URL `http://doi.acm.org/10.1145/1269900.1268790`.

[9] 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.

[10] 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.

[11] F.G. Lopez, R. W. Lent, S. D. Brown, and P. A. Gore. Role of social?cognitive expectations in high school students' mathematics-related interest and performance. *Journal of Counseling Psychology*, 44(1):44–52, 1997. doi: 10.1037/0022-0167.44.1.44.

[12] Milagros Sáinz and Jacquelynne Eccles. Self-concept of computer and math ability: Gender implications across time and within ict studies. *Journal of Vocational Behavior*, 80(2):486 – 499, 2012. ISSN 0001-8791. doi: https://doi.org/10.1016/j.jvb.2011.08.005. URL `http://www.sciencedirect.com/science/article/pii/S0001879111001126`.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] T. W. Price, J. Albert, V. Catete, and T. Barnes. Bjc in action: Comparison of student perceptions of a computer science principles course. In *2015 Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, pages 1–4, Aug 2015. doi: 10.1109/RESPECT.2015.7296506.

[19] Karen E Rambo-Hernandez Christina H Paguyo, Rebecca A Atadero and Jennifer Francis. Creating inclusive environments in first-year engineering classes to support student retention and learning. In *2015 ASEE Annual Conference & Exposition*, number 10.18260/p.23757, Seattle, Washington, June 2015. ASEE Conferences. https://peer.asee.org/23757.

[20] Jill Denner, Linda Werner, Lisa O'Connor, and Jill Glassman. Community college men and women: A test of three widely held beliefs about who pursues computer science. *Community College Review*, 42(4):342–362, 2014. doi: 10.1177/0091552114535624. URL `https://doi.org/10.1177/0091552114535624`.

[21] Lecia J. Barker, Charlie McDowell, and Kimberly Kalahar. Exploring factors that influence computer science introductory course students to persist in the major. *SIGCSE Bull.*, 41(1):153–157, March 2009. ISSN 0097-8418. doi: 10.1145/1539024.1508923. URL `http://doi.acm.org/10.1145/1539024.1508923`.

[22] Catherine T. Amelink and Elizabeth G. Creamer. Gender differences in elements of the undergraduate experience that influence satisfaction with the engineering major and the intent to pursue engineering as a career. *Journal of Engineering Education*, 99(1):81–92, 2010. ISSN 2168-9830. doi: 10.1002/j.2168-9830.2010.tb01044.x. URL `http://dx.doi.org/10.1002/j.2168-9830.2010.tb01044.x`.

[23] A.R. Brown, C. Morning, and C.B. Watkins. Implications of african american engineering student perceptions of campus climate factors. *Proceedings, 2004 ASEE/IEEE Frontiers in Education Conference*, 2004.

[24] A.R. Brown, C. Morning, and C.B. Watkins. Influence of african american engineering student perceptions of campus climate on graduation rates. *Journal of Engineering Education*, 94(2):263–271, 2005. doi: 10.1002/j.2168-9830.2005.tb00847.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2168-9830.2005.tb00847.x`.

[25] 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: Sample Survey Questions

How confident are you in your ability to be successful in this course?

	1	2	3	4	5	
Not at all confident	○	○	○	○	○	I can absolutely do this!

Do you think you were successful in this course? *

	1	2	3	4	5	
Not at all successful	○	○	○	○	○	Totally successful

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I find Computer Science intimidating.	○	○	○	○	○
I believe that other students in Computer Science will be welcoming of me.	○	○	○	○	○

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).

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	4.16	0.770	767
	URM	4.27	0.655	165
	Total	4.17	0.752	932
exitSurvey	non-URM	4.09	0.841	767
	URM	3.93	0.928	165
	Total	4.06	0.858	932

Mauchly's Test of Sphericity[a]

Measure:

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	11.257	1	11.257	27.141	0.000	0.028	27.141	0.999
	Greenhouse-Geisser	11.257	1.000	11.257	27.141	0.000	0.028	27.141	0.999
	Huynh-Feldt	11.257	1.000	11.257	27.141	0.000	0.028	27.141	0.999
	Lower-bound	11.257	1.000	11.257	27.141	0.000	0.028	27.141	0.999
time * URMStatus	Sphericity Assumed	5.008	1	5.008	12.075	0.001	0.013	12.075	0.935
	Greenhouse-Geisser	5.008	1.000	5.008	12.075	0.001	0.013	12.075	0.935
	Huynh-Feldt	5.008	1.000	5.008	12.075	0.001	0.013	12.075	0.935
	Lower-bound	5.008	1.000	5.008	12.075	0.001	0.013	12.075	0.935
Error(time)	Sphericity Assumed	385.734	930	0.415					
	Greenhouse-Geisser	385.734	930.000	0.415					
	Huynh-Feldt	385.734	930.000	0.415					
	Lower-bound	385.734	930.000	0.415					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:
Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	18342.100	1	18342.100	20768.581	0.000	0.957	20768.581	1.000
URMStatus	0.160	1	0.160	0.181	0.670	0.000	0.181	0.071
Error	821.344	930	0.883					

a. Computed using alpha = .05

Figure 7: Analysis for Course 1: Self-Efficacy. Within-subjects (time) is statistically significant, as is time * URM Status.
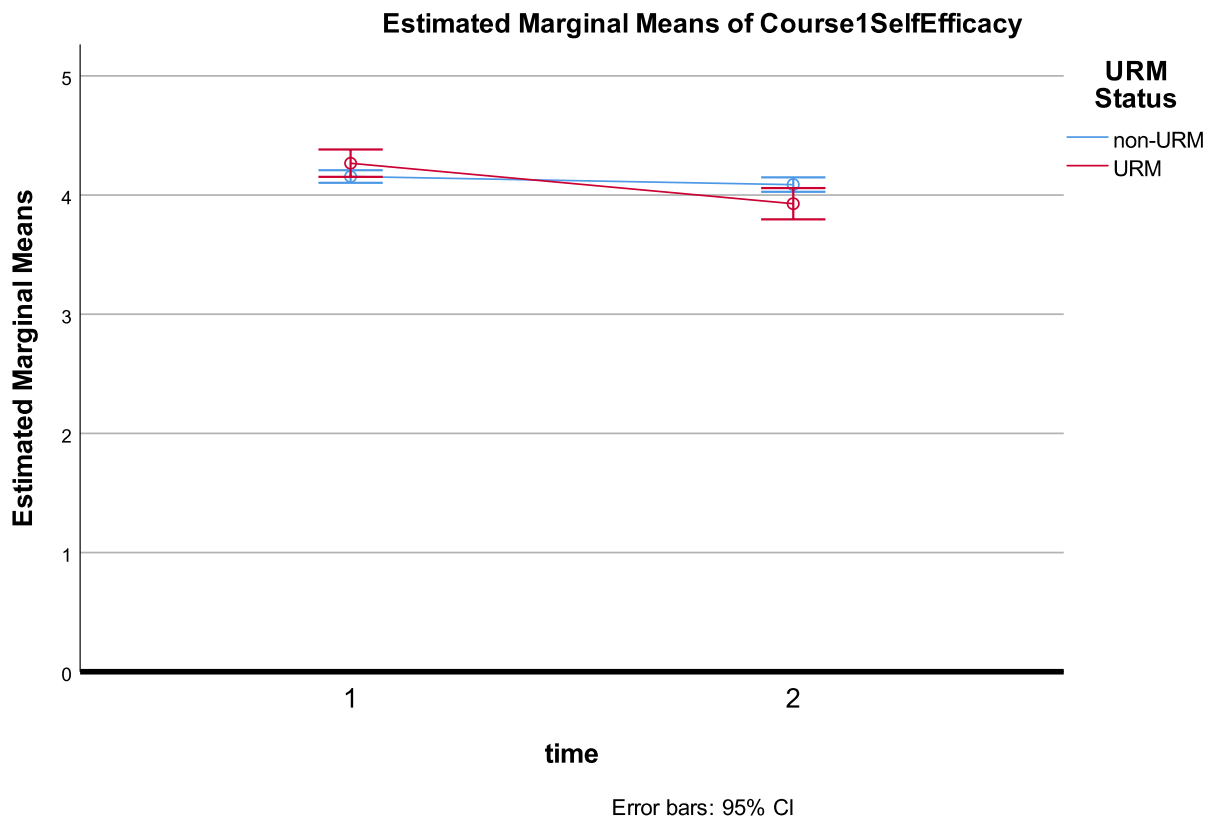
Figure 8: Plot for Course 1: Self-Efficacy. Within-subjects (time) is statistically significant; this means that the decrease in self-efficacy over the course of the term happens across the whole sample, although the size of the effect is small. Within-subjects (time * URM Status) is also statistically significant; this means that the decrease in self-efficacy is not equivalent across the two groups (URM and non-URM students). We note, however, that even with these slight decreases in self-efficacy, the mean responses are still well above the middle of the answer range – indicating that students in general still felt they were successful in the course.

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.18	1.169	767
	URM	3.08	1.099	165
	Total	3.16	1.157	932
exitSurvey	non-URM	3.11	1.229	767
	URM	3.18	1.189	165
	Total	3.12	1.222	932

Mauchly's Test of Sphericity[a]

Measure:

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	0.069	1	0.069	0.099	0.754	0.000	0.099	0.061
	Greenhouse-Geisser	0.069	1.000	0.069	0.099	0.754	0.000	0.099	0.061
	Huynh-Feldt	0.069	1.000	0.069	0.099	0.754	0.000	0.099	0.061
	Lower-bound	0.069	1.000	0.069	0.099	0.754	0.000	0.099	0.061
time * URMStatus	Sphericity Assumed	1.785	1	1.785	2.568	0.109	0.003	2.568	0.360
	Greenhouse-Geisser	1.785	1.000	1.785	2.568	0.109	0.003	2.568	0.360
	Huynh-Feldt	1.785	1.000	1.785	2.568	0.109	0.003	2.568	0.360
	Lower-bound	1.785	1.000	1.785	2.568	0.109	0.003	2.568	0.360
Error(time)	Sphericity Assumed	646.595	930	0.695					
	Greenhouse-Geisser	646.595	930.000	0.695					
	Huynh-Feldt	646.595	930.000	0.695					
	Lower-bound	646.595	930.000	0.695					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:

Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	10678.835	1	10678.835	4995.373	0.000	0.843	4995.373	1.000
URMStatus	0.071	1	0.071	0.033	0.856	0.000	0.033	0.054
Error	1988.103	930	2.138					

a. Computed using alpha = .05

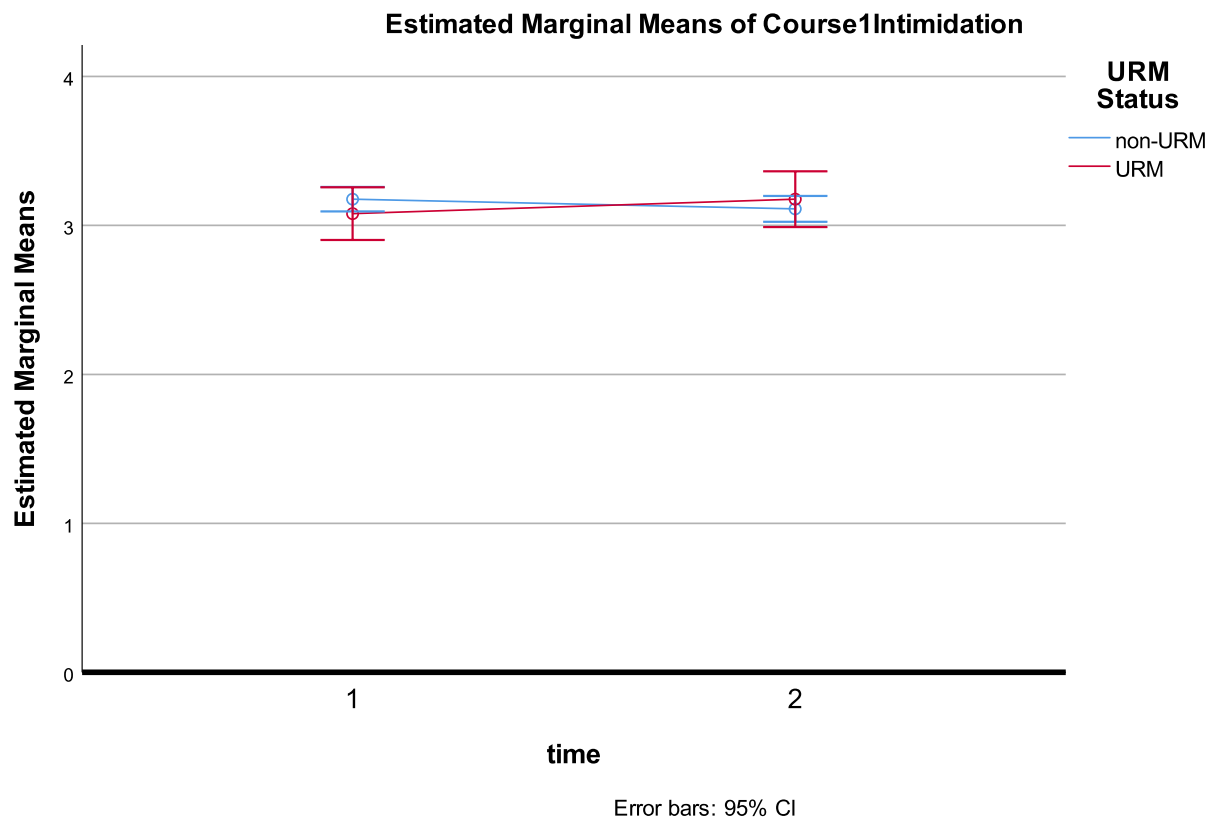Figure 9: Analysis for Course 1: Intimidation. No statistical significance.

Figure 10: Plot for Course 1: Intimidation. No statistical significance.

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.81	0.783	767
	URM	3.80	0.759	165
	Total	3.81	0.779	932
exitSurvey	non-URM	3.75	0.909	767
	URM	3.62	0.972	165
	Total	3.73	0.921	932

Mauchly's Test of Sphericity[a]

Measure:

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	3.884	1	3.884	8.420	0.004	0.009	8.420	0.826
	Greenhouse-Geisser	3.884	1.000	3.884	8.420	0.004	0.009	8.420	0.826
	Huynh-Feldt	3.884	1.000	3.884	8.420	0.004	0.009	8.420	0.826
	Lower-bound	3.884	1.000	3.884	8.420	0.004	0.009	8.420	0.826
time * URMStatus	Sphericity Assumed	1.052	1	1.052	2.280	0.131	0.002	2.280	0.326
	Greenhouse-Geisser	1.052	1.000	1.052	2.280	0.131	0.002	2.280	0.326
	Huynh-Feldt	1.052	1.000	1.052	2.280	0.131	0.002	2.280	0.326
	Lower-bound	1.052	1.000	1.052	2.280	0.131	0.002	2.280	0.326
Error(time)	Sphericity Assumed	429.011	930	0.461					
	Greenhouse-Geisser	429.011	930.000	0.461					
	Huynh-Feldt	429.011	930.000	0.461					
	Lower-bound	429.011	930.000	0.461					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:
Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	15235.738	1	15235.738	15343.357	0.000	0.943	15343.357	1.000
URMStatus	1.403	1	1.403	1.413	0.235	0.002	1.413	0.221
Error	923.477	930	0.993					

a. Computed using alpha = .05

Figure 11: Analysis for Course 1: Inclusion. Within-subjects (time) is statistically significant.

Estimated Marginal Means of Course1Inclusion
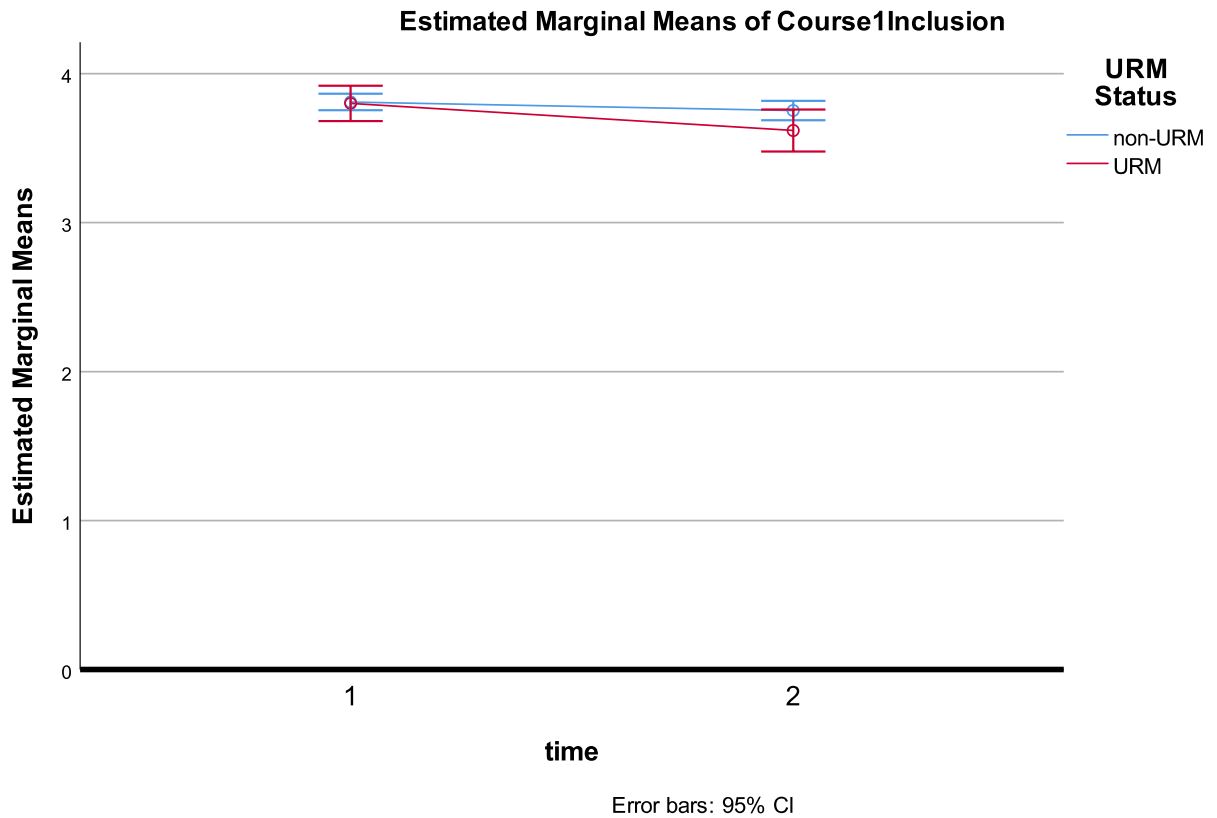
Error bars: 95% CI

Figure 12: Plot for Course 1: Inclusion. Within-subjects (time) is statistically significant. This means that the decrease in feeling welcome over the course of the term happens across the whole sample, although the size of the effect is small. The decrease in feeling welcome is a little larger for URM students, but it is not statistically significant.

Descriptive Statistics

	URM Status	Mean	Std. Deviation	N
entrySurvey	non-URM	3.92	0.860	1558
	URM	3.93	0.837	208
	Total	3.92	0.857	1766
exitSurvey	non-URM	3.76	0.873	1558
	URM	3.73	0.925	208
	Total	3.76	0.879	1766

Mauchly's Test of Sphericity[a]

Measure:

					Epsilon[b]		
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	11.837	1	11.837	23.945	0.000	0.013	23.945	0.998
	Greenhouse-Geisser	11.837	1.000	11.837	23.945	0.000	0.013	23.945	0.998
	Huynh-Feldt	11.837	1.000	11.837	23.945	0.000	0.013	23.945	0.998
	Lower-bound	11.837	1.000	11.837	23.945	0.000	0.013	23.945	0.998
time * URMStatus	Sphericity Assumed	0.183	1	0.183	0.370	0.543	0.000	0.370	0.093
	Greenhouse-Geisser	0.183	1.000	0.183	0.370	0.543	0.000	0.370	0.093
	Huynh-Feldt	0.183	1.000	0.183	0.370	0.543	0.000	0.370	0.093
	Lower-bound	0.183	1.000	0.183	0.370	0.543	0.000	0.370	0.093
Error(time)	Sphericity Assumed	871.996	1764	0.494					
	Greenhouse-Geisser	871.996	1764.000	0.494					
	Huynh-Feldt	871.996	1764.000	0.494					
	Lower-bound	871.996	1764.000	0.494					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:
Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	21599.356	1	21599.356	21287.874	0.000	0.923	21287.874	1.000
URMStatus	0.024	1	0.024	0.024	0.877	0.000	0.024	0.053
Error	1789.811	1764	1.015					

a. Computed using alpha = .05

Figure 13: Analysis for Course 2: Self-Efficacy. Within-subjects (time) is statistically significant.

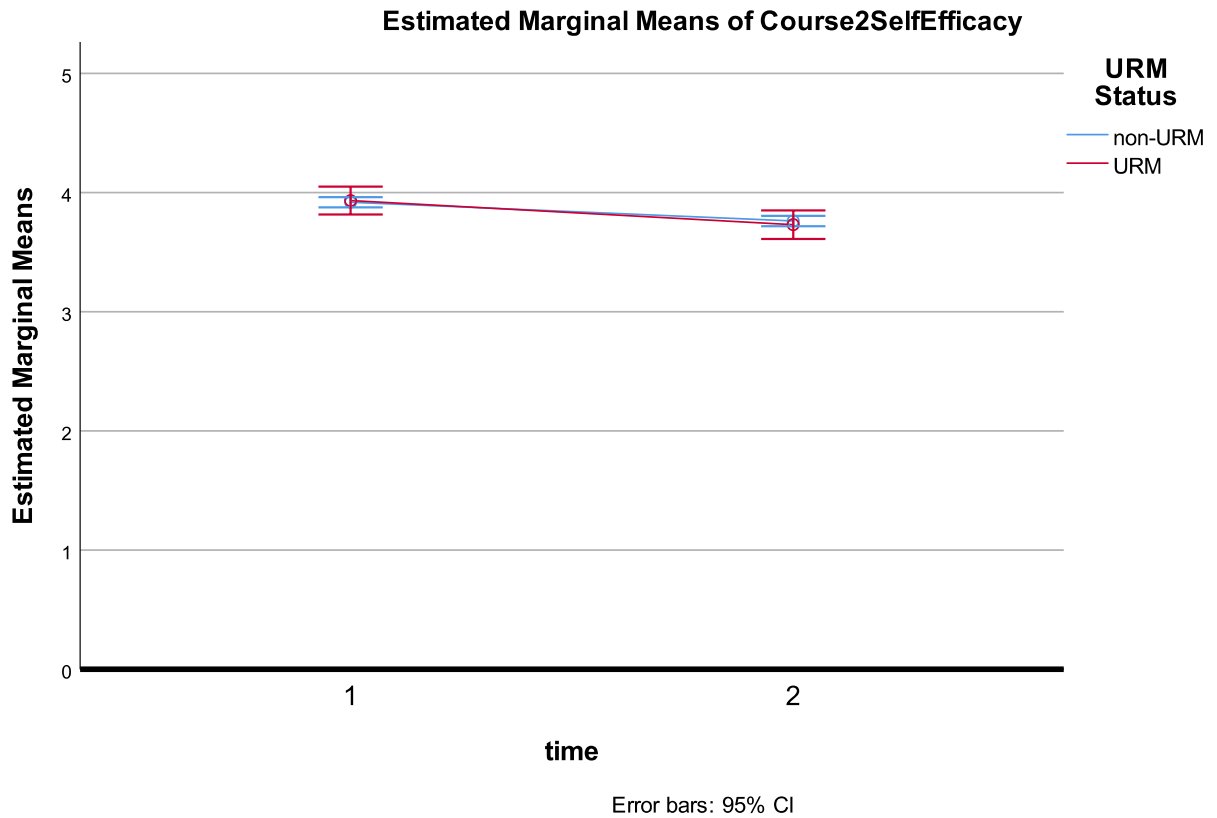**Estimated Marginal Means of Course2SelfEfficacy**

Error bars: 95% CI

Figure 14: Plot for Course 2: Self-Efficacy. Within-subjects (time) is statistically significant. This means that the decrease in self-efficacy over the course of the term happens across the whole sample, although the size of the effect is small. We note that there appears to be no difference between URM and non-URM students' reported self-efficacy for this course, both statistically and visually.

Descriptive Statistics

	URM Status	Mean	Std. Deviation	N
entrySurvey	non-URM	3.07	1.184	1558
	URM	3.02	1.110	208
	Total	3.06	1.175	1766
exitSurvey	non-URM	3.12	1.200	1558
	URM	3.05	1.160	208
	Total	3.11	1.195	1766

Mauchly's Test of Sphericity[a]

Measure:

					Epsilon[b]		
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	0.619	1	0.619	0.930	0.335	0.001	0.930	0.161
	Greenhouse-Geisser	0.619	1.000	0.619	0.930	0.335	0.001	0.930	0.161
	Huynh-Feldt	0.619	1.000	0.619	0.930	0.335	0.001	0.930	0.161
	Lower-bound	0.619	1.000	0.619	0.930	0.335	0.001	0.930	0.161
time * URMStatus	Sphericity Assumed	0.055	1	0.055	0.082	0.774	0.000	0.082	0.059
	Greenhouse-Geisser	0.055	1.000	0.055	0.082	0.774	0.000	0.082	0.059
	Huynh-Feldt	0.055	1.000	0.055	0.082	0.774	0.000	0.082	0.059
	Lower-bound	0.055	1.000	0.055	0.082	0.774	0.000	0.082	0.059
Error(time)	Sphericity Assumed	1173.203	1764	0.665					
	Greenhouse-Geisser	1173.203	1764.000	0.665					
	Huynh-Feldt	1173.203	1764.000	0.665					
	Lower-bound	1173.203	1764.000	0.665					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:

Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	13799.177	1	13799.177	6436.707	0.000	0.785	6436.707	1.000
URMStatus	1.107	1	1.107	0.516	0.472	0.000	0.516	0.111
Error	3781.708	1764	2.144					

a. Computed using alpha = .05

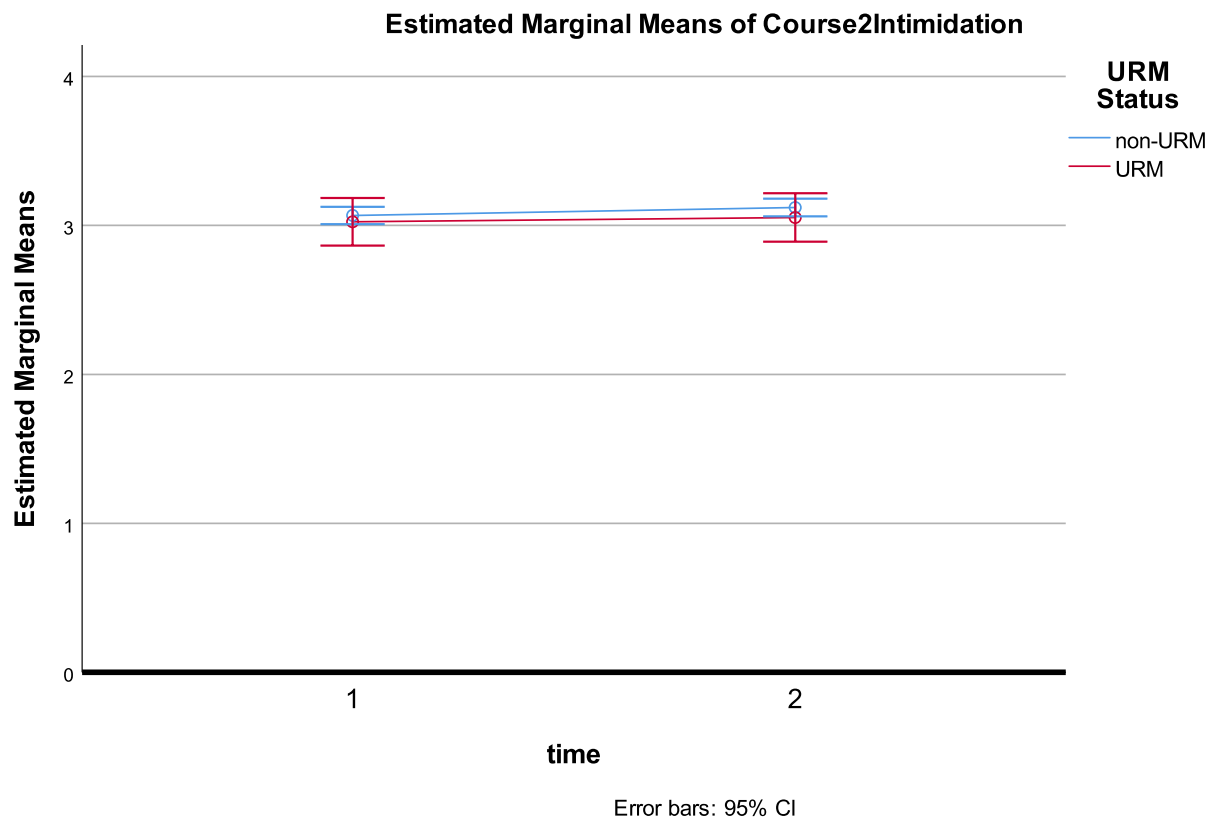Figure 15: Analysis for Course 2: Intimidation. No statistical significance.

Figure 16: Plot for Course 2: Intimidation. No statistical significance.

Descriptive Statistics

URM Status		Mean	Std. Deviation	N
entrySurvey	non-URM	3.89	0.809	1558
	URM	3.79	0.879	208
	Total	3.88	0.818	1766
exitSurvey	non-URM	3.84	0.899	1558
	URM	3.71	0.965	208
	Total	3.82	0.908	1766

Mauchly's Test of Sphericity[a]

Measure:

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
						Epsilon[b]	
time	1.000	0.000	0		1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + URMStatus
 Within Subjects Design: time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Tests of Within-Subjects Effects

Measure:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
time	Sphericity Assumed	1.547	1	1.547	3.936	0.047	0.002	3.936	0.509
	Greenhouse-Geisser	1.547	1.000	1.547	3.936	0.047	0.002	3.936	0.509
	Huynh-Feldt	1.547	1.000	1.547	3.936	0.047	0.002	3.936	0.509
	Lower-bound	1.547	1.000	1.547	3.936	0.047	0.002	3.936	0.509
time * URMStatus	Sphericity Assumed	0.104	1	0.104	0.263	0.608	0.000	0.263	0.081
	Greenhouse-Geisser	0.104	1.000	0.104	0.263	0.608	0.000	0.263	0.081
	Huynh-Feldt	0.104	1.000	0.104	0.263	0.608	0.000	0.263	0.081
	Lower-bound	0.104	1.000	0.104	0.263	0.608	0.000	0.263	0.081
Error(time)	Sphericity Assumed	693.500	1764	0.393					
	Greenhouse-Geisser	693.500	1764.000	0.393					
	Huynh-Feldt	693.500	1764.000	0.393					
	Lower-bound	693.500	1764.000	0.393					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:
Transformed Variable:

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	21280.352	1	21280.352	19375.387	0.000	0.917	19375.387	1.000
URMStatus	4.434	1	4.434	4.037	0.045	0.002	4.037	0.519
Error	1937.434	1764	1.098					

a. Computed using alpha = .05

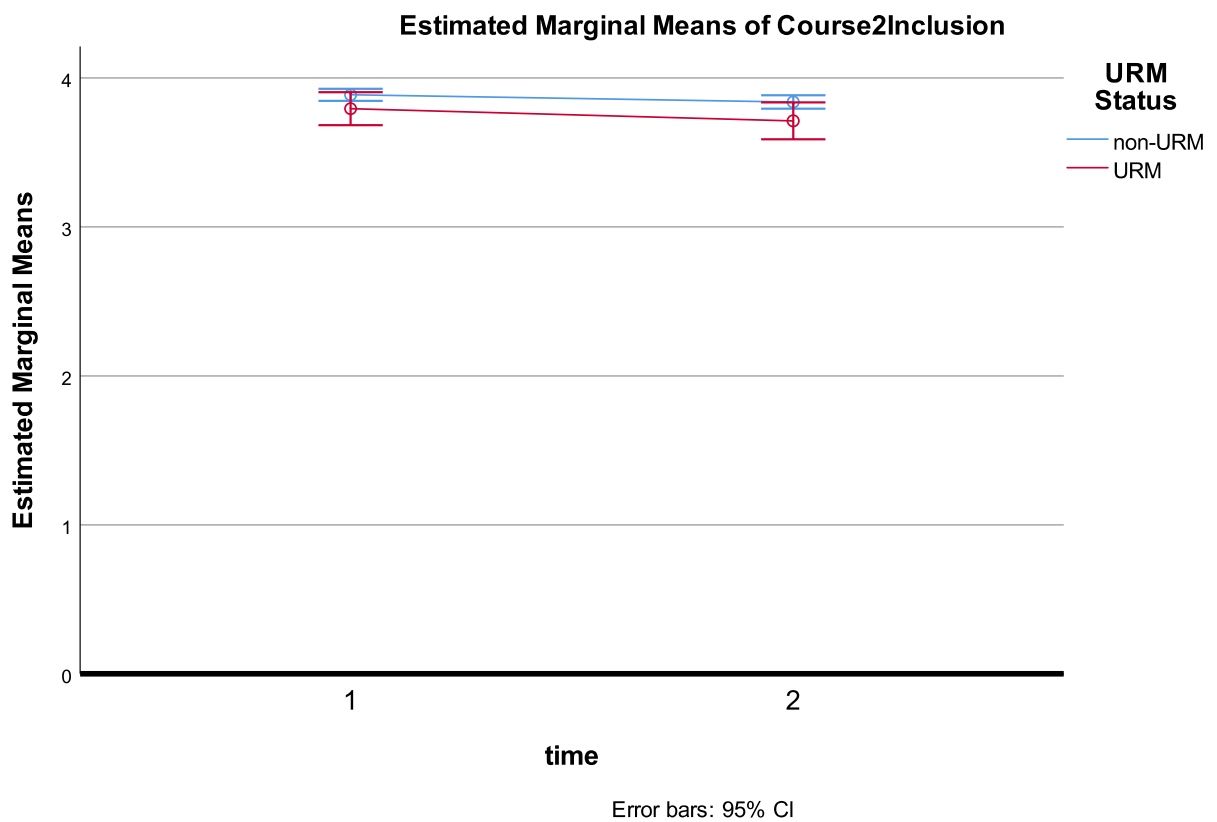Figure 17: Analysis for Course 2: Inclusion. No statistical significance.

Figure 18: Plot for Course 2: Inclusion. No statistical significance. Visually, we note that URM students report feeling less welcome, and even though it was not statistically significant in this data set, we should be cognizant of this difference in perception.