Understanding Underrepresentation in Computer Science at Portland State University

Danielle Beyer
Computer Science
Portland State University
Portland, Oregon, USA
Email: wishnuff@pdx.edu

Sarah Haskew
Computer Science
Portland State University
Portland, Oregon, USA
Email: sahaskew@pdx.edu

Alyssa Tamayo
Computer Science
Portland State University
Portland, Oregon, USA
Email: atamayo@pdx.edu

ABSTRACT
As technology becomes more integrated in everyday life, software needs to represent everyone's experiences. Software design decisions impact the way we live our lives, so the diversity of those designing software is becoming more important. This study analyzes data regarding retention and attrition by gender in the computer science department at Portland State University. We offer a look into the gender disparity problem in computer science by studying the history of recruiting women and understanding the challenges that come with retrieving retention data. The next stage of this study is to interview student experiences.

CCS CONCEPTS
•Social and professional topics -->Professional topics ~Computing education ~Computing education programs ~Computer science education ~CS1 •General and reference -->Document type s~Surveys and overviews •General and reference ~Document types~Reference works •Social and professional topics -->Professional topics ~Computing industry ~Industry statistics •Social and professional topics -->Professional topics ~Computing education ~K-12 education

KEYWORDS
diversity; inclusion; equity; women; underrepresented groups; education; university; computer science; recruitment; retention; persistence; attrition; belonging; interviews;

ACM Reference format:
D. Beyer, S. Haskew and A. Tamayo. 2020. Understanding Underrepresentation in Computer Science at Portland State University. In *Proceedings of CS 410/510: Explorations of Data Science, Portland, OR, USA, 10 pages.*

1. Introduction

As computer science students at Portland State University (PSU), we noticed a lack of diversity in our lower-division classes which only became more noticeable as we progressed through the curriculum. Our perception of this lack of diversity in our collegiate setting is echoed in other post-secondary academic institutions as well as the industry as a whole.

The lack of diversity in Science, Technology, Engineering and Math (STEM) related fields is problematic. Groups which are underrepresented in academia make up less of the jobs in STEM fields than they do in the overall workforce, for example women hold 57% of professional occupations while in science and engineering they represent only 28% of the workforce. [1] These percentages are lower when you look at computing and technology fields.

In this paper, we offer our analysis of PSU CS diversity data and the greater US CS diversity data. We present existing research regarding challenges in obtaining retention data as well as compare our data collection experience to previous efforts. Though our intent is to examine data for all underrepresented groups at PSU, not just women, we also study historical efforts in recruiting women. Unfortunately, our existing dataset only compares women to the overall PSU CS student body and we hope to obtain data with more demographic detail including racial identity, age and grades. Our future work also includes collecting experiences of students in the PSU CS department.[1]

2. Motivation

Underrepresentation of diverse groups within computer science (CS) majors in the United States impacts the ongoing workforce disparity. Our intent is to examine how

[1] While we reference women here, we understand that not all students identify are male or female identifying. At this time, we are limited to data which represents only "legal sex".

PSU's CS department compares to the overall US state of CS education.

Recently, the PSU Maseeh College of Engineering & Computer Science (MCECS) celebrated the earnings for recent graduates of each of their majors, with CS majors earning the more than any other engineering related major by at least 15%. [2] While laudable, this also makes the issue of lack of representation in the CS department even more dire by limiting access to social mobility for the groups which are underrepresented in CS education.

As technology becomes more integrated in everyday life, software needs to represent everyone's experiences. Software design decisions impact the way we live our lives, so the diversity of those designing software is becoming more important. If one type/background person is creating software, it might be creating software that only serves the demographic creating it. The needs of the people using software are not represented by those building the software.

3. Data

3.1 PSU Existing Data

James Hook, Associate Dean, MCECS, had previously presented on the retention of women in the CS department and kindly shared his findings with us. We received dashboards from him that grouped students by academic standing which included freshmen through seniors that have not yet been admitted into the CS department, admitted juniors and seniors, pre-admitted and admitted post baccalaureate students and graduate students. Two main dashboards tracked student count from 2014 to 2019; one containing all students regardless of gender and another tracking only women. We used these existing dashboards to calculate that only 11% of CS undergraduate degrees awarded were awarded to women. [3] For reference, the national average is 19% of all CS undergraduate degrees are awarded to women [4].[2]

3.2 PSU Data Request

Though Hook's data gave us a brief look into the diversity of our department, it was evident that we needed to probe for more data elsewhere.

With Hook's help, we were able to make a connection to the Office of Institutional Research and Planning (OIRP). The OIRP will help us retrieve more detailed data in regards to gender, race/ethnicity and other important groups. We were presented with two problems: as a response to departmental loss of income due to COVID-19, faculty were furloughed on Fridays limiting any possible free-time the team would have to work on these queries on our behalf. Second, the data is not de-identified and any efforts to do so that might allow us access to the database would be too costly. We will be limited to only the queries the OIRP will do on our behalf subject to the data set being large enough to warrant student privacy. Thus, the third unfortunate truth, queries relating to underrepresented students, black students especially, at our department might turn up counts too small triggering privacy concerns. This necessitates the aggregation of data over a larger period of time and limits our ability to analyze trends year over year.

3.3 Data Access Beyond PSU

The problems we have encountered in accessing data are not unique. In the Association for Computing Machinery (ACM) study *Retention in Computer Science Undergraduate Programs in the U.S. – Data Challenges and Promising Interventions* from 2018, the ACM echoes our experience. [5] The report echoes two of our key data challenges, the inability to track an individual's progress through their course of study and the necessity of looking at aggregate data without being able to identify patterns for race/ethnicity and gender.

The ACM provides two case studies in which data was requested from universities, the University of California, San Diego (UCSD) and Colorado School of Mines. While each case was unique, the researchers had common challenges which we have encountered as well.

- Database access was limited to defined data analysts
- The database is complicated, requiring subject matter expertise in data fields and their meaning
- Structuring a data request that was meaningful to the data analyst was challenging
- It took more than a month to actually receive data (at UCSD it took multiple months)
- Students have multiple entry points to the university and CS program and defining how and when to include them into the dataset was not straightforward

Ultimately, for both case studies, ACM was able to acquire a full data set that included data for each individual student that was able to both anonymize the data and still allow tracking of the individual student's progress. We have been told that this will not be possible and that if we can acquire the data we have requested, that it will be aggregated.

4. Recruiting Women

In 1967 Grace Hopper was quoted in a *Cosmopolitan Magazine* article entitled "The Computer Girls". Talking about the field of computer programming, she said "It's just like planning a dinner. You have to plan ahead and schedule everything so it's ready when you need it. Programming requires patience and the ability to handle detail**.** Women are naturals at computer programming." [6] Today's modern audience will roll their eyes on the first

[2] Due to the limitations of Hook's original data source, the data itself didn't track students who identified outside of the binary gender norm. The dashboards were forced to make trivial assumptions about gender identity. We hope to be able to address this with future data pulls.

statement and not get to the second. When Dr. Hopper said this, women were actively recruited as computer programmers with little to no experience by taking an aptitude test which tested for logic. Considering the era, Stanford had only just created a Computer Science department two years earlier. Coding jobs were exploding as corporations began rolling out automated processes and employers looked for logical, meticulous candidates while gender stereotypes supported the candidacy of women. Women continued to gain share of the programming industry until 1984. [7]

Since 1984, representation of women in both computer science departments and in industry has declined, while women's representation in other challenging fields has increased. [8] What has driven this decline and why are women not choosing to pursue computer science?

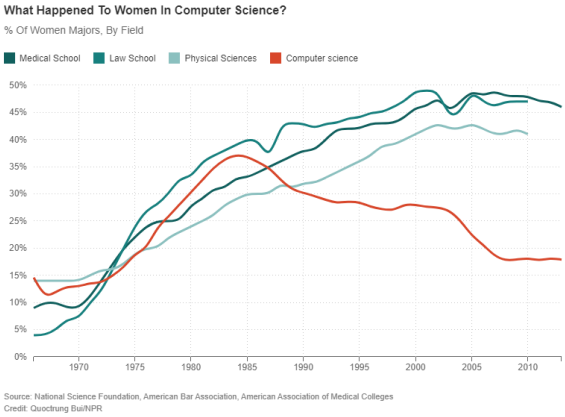

Figure 1: Women's Representation in Selected Professional Fields of Study [8]

One theory that is reiterated across the published literature is that the rise of the personal computer initiated the decline of women in computer science. As PCs became available, they were marketed to boys. They were purchased by families for their sons. As gaming became more popular, this only became more entrenched in the marketing gender imbalance.

One anecdote shared by NPR [8] is that of a mathematically talented female student who dropped out of Johns Hopkins University's CS program after her first introductory course. She shared that the men in her class came in with a level of understanding she didn't have and that despite the introductory nature of the course, the instructor belittled those who hadn't come into the class with the basics. It took her ten years but she did eventually earn a Ph.D. in computer science.

Her story is echoed in many others throughout the literature. Prior to the boom of the PC, instructors expected that students came into their classes with no prior knowledge. As student experience began changing in the 1980s, professors changed their expectations as well – leaving those without that experience at a disadvantage.

As women began to disappear from CS classes, the lack of gender diversity became self-perpetuating. Despite the changing commercial environment for computers – they are now ubiquitous -- women didn't see themselves portrayed in the profession nor did they see themselves in CS classes, they didn't see those as goals for them.

Despite the earning potential of a CS degree, today, Computer Science doesn't even make the list of the top 20 most preferred majors by women. While men's list of 20 most preferred majors includes Computer Science and Computer Information Systems.

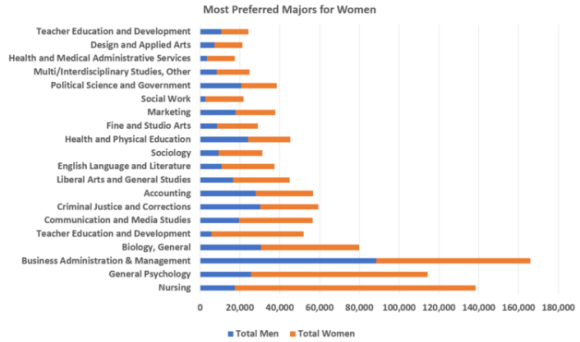

Figure 2: Women's 20 Most Preferred Majors [9]

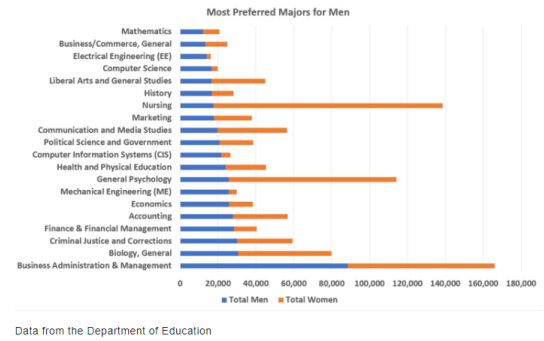

Figure 3:Men's 20 Most Preferred Majors [9]

So, what can be done about this? There are various recommendations and outreach efforts underway. The National Center for Women & Information Technology (NCWIT) has many published programs to support recruitment and engagement at the K-12 level. [4] The College Board has recruitment programs to improve the representation of women taking the CS Advanced Placement exams. They have even instituted an annual AP Computer Science Female Diversity Award to schools that have achieved either 50% or higher female representation in the AP exam or which the percentage meets or exceeds the female percentage of the school population. [10]

As we continue to examine this topic, we will further expand this section on efforts to improve recruitment of women into CS studies.

5. Retaining Women

While it's clear that there is a recruitment problem in getting women to choose CS, the problem transitions to a retention or persistence problem once students have begun taking CS classes. At PSU, only 13% of students in lower division classes were women in 2016. Of those, 43% did not persist within CS to the 2017-2018 academic year. [3] Even if we can recruit women into the department, PSU is losing them before they attain a CS degree.

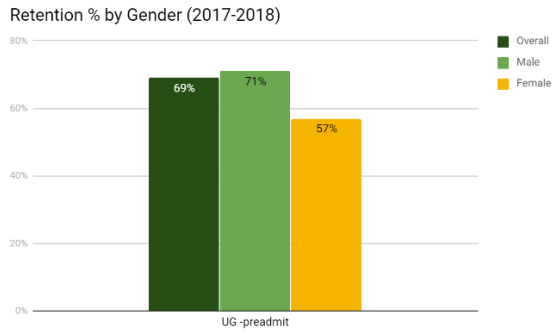

Figure 4: PSU Lower Division Retention by Gender [3]

PSU's numbers are far worse than overall in the US. Overall retention for Freshmen women in the US in CS was 73% in 2014-15 [5]. UCLA's Senior Researcher, Dr. Jane Margolis has pioneered the study of retention for underrepresented groups in CS. Her research, begun in 1994, sought to break down the problem of retention and understand solutions for each cause. Margolis interviewed CS students at Carnegie Mellon (CMU) over a four-year period, ultimately publishing her findings in her book "Unlocking the Clubhouse: Women in Computing" and making suggestions that were implemented by the CMU CS department. [7] CMU has had about a 50% class split of men vs. women and an 89% retention rate for several years. [5]

What did CMU do that is different? They defined five critical conditions, which are described in detail in the ACM's report on diversity and retention. [5]

- Institutional commitment and multiple levels of support
- Assessment of context and environment
- Ongoing attention paid to the situation
- A community that is open to cultural change
- Assurance that women are central to the culture

Stated less formally, they removed the barriers to entry and ensured that the classroom and university environment supported their students. CMU removed requirements for prior coding experience. They implemented various tracks to get through the lower division coding required coursework based on the amount of prior coding experience. They implemented a peer mentoring and professional mentoring program. And, they held their faculty accountable for ensuring all students were supported and welcomed, not just those who came in with enough knowledge to impress their professors.

Similarly, Harvey Mudd College (HMC) made changes to their CS program and has been lauded with accolades for the diversity in their CS graduates. [11] HMC offers multiple tracks for its introductory course based on experience. They offer three different versions of their first introductory CS course. A fourth version, in which the first and second introductory courses are combined, is offered for those coming in with a programming background. The goal is that all students finishing the first course are prepared for the second course; that all students completing the second course are prepared to move on in CS. They insist that professors set the expectation that hard work and spending time getting help are the keys to success, not getting the right answer. Professors are tasked with creating a welcoming environment and keeping the more intimidating students from dominating the classroom. Finally, students are strongly encouraged to complete the two-class series, at the end of which students ideally have the confidence to continue on.

While both CMU and HMC are successful, they have very different environments and student bodies. Elements of their success can be leveraged and we hope to incorporate their successes into our ultimate findings.

With our interviews, we hope to further understand how PSU students view their experiences and how their experiences in the classroom and on campus have impacted their decisions to continue in CS or pursue something else. Margolis' studies continue at UCLA, and we intend to leverage her studies as we evaluate our own research outcomes. [12]

6. Interview Approval

Recognizing that the PSU CS department has a high attrition rate -- especially for students who are members of groups historically underrepresented in computing -- we seek to investigate why some students persist and why other students choose to leave the program. To further this investigation, we intend to conduct semi-structured qualitative interviews of current and former students regarding their experiences in the CS program at PSU in order to better understand persistence, attrition, retention, and success.

Interviews are categorized as "Human Research". Prior to conducting Human Research, approval must be acquired from the Review Board (IRB). The IRB submission includes a script for recruiting and a set of questions to ask potential participants. The submitted script is based on what we thought would draw both people who felt excluded from Computer Science and those who felt included. Though our aim is to hear stories from those who experienced discrimination or felt othered to the point of dropping out, we also want to hear from students who stayed in the program, even if they had considered leaving.

Analysis of both experiences will provide a richer basis for reasons for attrition and retention.

The IRB documentation was submitted and approved. We have been approved to move forward with the interview process, further described in section *6. Future Work*.

7. Future Work

At the time of writing, we have completed an initial analysis of existing data and requested data with greater detail for future analysis. We have obtained approval to begin our qualitative study of student lived experiences. We have the following phases of our qualitative study: recruitment, interviews, and analysis.

7.1 Interviews

7.1.1 Recruitment

Recruitment for interviews has begun and will continue through email campaigns to students, postings on various social networks (including the department's slack channel, personal Facebook pages, personal Instagram accounts, etc.) as well as, leveraging personal connections with current and previous students in our department. We will leverage a snowball sampling technique to recruit further participants. We hope to interview the following people regardless of their gender identity, country of residence, race/ethnicity, or sexual orientation: current PSU CS students, graduated PSU CS students and students who left the CS program. This is important because different levels of the program have different retention rates and various experiences. The IRB approval process included approval of our recruitment emails, graphics and social network posts.

7.1.2 Interviews

Interviews are semi-structured [13], meaning that while we have prepared an order of categories and a set of questions, each interview will be adapted based on the study progression and individual participants. Because we will have multiple interviewers, we expect to stay relatively close to an outline of planned questions. The guide questions are grouped into three categories, background information, motivation and interest, adversity and closing.

Background information will be used to give context and compare the experiences from other interviews. This section will collect basic personal information such as the interviewee's age, gender and sexual orientation. This section will also explore any previous education or work experience prior to their studies at PSU's CS department as well as get a timeline for which CS related classes they took in our department or elsewhere (i.e. community college, coding bootcamps, self-taught, etc.)

While the next section is named *Motivation and Interest*, this section also explores the interviewee's support in the program. Here we probe for any types of support they have access to while pursuing their studies in CS. We open up the questioning here to be free-flowing and to allow the interviewee to share anything about their personal experience thus far to the interviewers. We aim to probe for what motivates them to continue or discontinue their studies in computer science.

Adversity questions may be modified as the study progresses or tailored to individual respondents as appropriate, but will always focus on discussing adversity, challenges, & barriers. Here we discuss any discrimination or microaggression they may have faced from a fellow student or faculty member. We hope to get a sense of whether this student has or dropped out, or thought about dropping out, of the program due to lack of support at home or within the department.

7.1.3 Analysis of Interviews

We intend to conduct interviews remotely over Zoom. Calls will be recorded in order to allow for transcription. Interviews will be transcribed and all identifying information will be omitted from the transcription and the analysis. Once transcribed, the recording will be deleted. We plan to use the software *Atlasti* to organize our data and analyze for trends.

7.2 Outcomes

After analyzing both the interviews and any additional data we are able to obtain from OIRP, the end result is to make our findings accessible to students and faculty in our department. Here we aim to share insights on diversity statistics within the PSU CS department and to amplify the voices of underrepresented students in the department while offering valuable insight to the department on how they can better support these students.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Kristin Tufte for her overarching support and sponsorship of this project, Dr. Ellie Harmon for her guidance on the IRB process and her sponsorship on this project and Dr. James Hook for his help in accessing and understanding data about diversity within the PSU CS department.

References

[1] "NGC Project," March 2018. [Online]. Available: https://ngcproject.org/sites/default/files/ngcp_the_state_of_girls_and_women_in_stem_2018a.pdf. [Accessed 14 August 2020].

[2] "Maseeh College of Engineering & Computer Science," Portland State University, July 2020. [Online]. Available: https://www.pdx.edu/cecs/the-maseeh-college-of-engineering-and-computer-science-transforms-lives. [Accessed 14 August 2020].

[3] J. Hook, Interviewee, *Associate Dean, MCECS, PSU.* [Interview]. July 2020.

[4] "NCWIT," [Online]. Available: www.ncwit.org. [Accessed 13 August 2020].

[5] Association for Computing Machinery (ACM), "Retention in Computer Science Undergraduate Programs in the U.S. -- Data

Challenges and Promising Interventions," Association for Computing Machinery, New York, 2018.

[6] L. Mandel, "The Computer Girls," *Cosmopolitan,* pp. 52-56, April 1967.

[7] C. Thompson, "New York Times Magazine," 13 February 2019. [Online]. Available: https://www.nytimes.com/2019/02/13/magazine/women-coding-computer-programming.html. [Accessed 11 August 2020].

[8] S. Henn, "National Public Radio," 21 October 2014. [Online]. Available: https://www.npr.org/sections/money/2014/10/21/357629765/when-women-stopped-coding. [Accessed 12 August 2020].

[9] "College Factual," 22 November 2017. [Online]. Available: https://inside.collegefactual.com/stories/the-most-popular-majors-for-women-men. [Accessed 8 August 2020].

[10] College Board, "College Board," 27 January 2020. [Online]. Available: https://www.collegeboard.org/releases/2020/20-more-schools-earn-annual-ap-computer-science-female-diversity-award. [Accessed 13 August 2020].

[11] K. Weisul, "Half of This College's STEM Graduates Are Women. Here's What They Did Differently," 31 May 2017. [Online]. Available: https://www.inc.com/kimberly-weisul/how-harvey-mudd-college-achieved-gender-parity-computer-science-engineering-physics.html. [Accessed 14 August 2020].

[12] S. Vu, "UCLA," 26 June 2017. [Online]. Available: https://newsroom.ucla.edu/stories/cracking-the-code:-why-aren-t-more-women-majoring-in-computer-science. [Accessed 13 August 2020].

[13] H. R. Bernard, "Interviewing: Unstructured & Structured," in *Research Methods in Anthropology: Qualitative & Quantitative Approaches, 4h Ed.* , Altamira Press, 2006.

APPENDIX

1. Interview Guide

Interview Guide

Interviews will be semi-structured[1] and adapted based on study progression and individual participants[2]. An outline of planned questions appears below. Because we have multiple interviewers, we expect to stay relatively close to the guide. See further notes in each section.

Start the interview by going over the consent form, including confidentiality assurance, purpose of project, participant questions. Verify again that the interview will be recorded, but file deleted after transcription.

Background information

Note: *All participants will be asked these basic survey questions which will not change through the course of the study*

To participant: I want to start with some background information to help me put your answers in context with other interviews. Remember, none of this information will be shared outside the project.

1. How old are you?
2. Please describe your gender identity.
3. Please describe your racial identity.
4. If comfortable, please describe your sexuality.
5. What is the highest educational attainment for each of your parents?
    a. *Choices for each parent if needed: some high school, high school, some college, 2-year degree, 4-year degree, some graduate school, graduate degree*
6. What was your educational background before you entered PSU?
    a. *Followup probes for both education level (e.g., prior BS, some college, GED, high school diploma) and specific to CS/programming (e.g., any experience in a coding camp as a high schooler).*
7. So I want to just sketch out a high-level timeline of your PSU experience, when you started, any times you've taken a break, and where you're at now [*if former student, when you left graduated / if current student, when you expect to graduate*].
    a. Probe for work: full time student vs. part time or full time worker in addition to school, has employment changed through PSU experience
    b. *Goal to capture a brief, high-level overview, e.g.*, attended from 1999-2001 part-time while also working, left for a while to work full time, back from 2005-2008 as a full-time student, or graduated w/ BS in 2008.
8. I have a list here of all the core CS courses, I just want to go through them with you real quick, and have you tell me whether you took these courses at PSU or elsewhere, and if you ever had to retake any classes. I don't need the specific grade you got in each class. (See table on last page for check off sheet to be used).

[1] Bernard, H. Russel. "Interviewing: Unstructured & Structured." In *Research Methods in Anthropology: Qualitative & Quantitative Approaches*, 4th ed. Altamira Press, 2006.
[2] Corbin, Juliet M, and Anselm C Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 3rd ed. Sage Publications, Inc, 2007.

Motivation & Interest

Note: Specific questions in this section may be modified as the study progresses and tailored to individual respondents as appropriate, but will always focus on general ***motivation, support, & interest*** *in CS.*

9. OK, so shifting gears from all those basic questions, I'd like to hear more about your personal story. To start, can you tell me about when you first decided to pursue a degree in computer science.
    a. *Probe for: initial motivation / interest, when this was relative to present day, was this an idea before starting college, or the result of some specific event, a friend who has this degree, etc.*
10. Tell me about the sources of support you have for pursuing your degree.
    a. *Probe for: family, friends, student groups, scholarships / financial aid, mentor*
11. Tell me about what motivates you to continue your studies? What helps you keep going!
12. Tell me about a time when you felt like you "belonged" in CS or felt like your contribution to CS was valued?
    a. *Probe for: More examples? How often does this happen? Is it rare or frequent?*

Adversity

Note: Specific questions in this section may be modified as the study progresses or tailored to individual respondents as appropriate, but will always focus on discussing ***adversity, challenges, & barriers****.*

13. And now, switching gears a little again, can you tell me about a time when you felt like you *didn't* belong or were *not* valued?
    a. *Probe for more examples, frequency*
14. Tell me about a really challenging or difficult experience you've had as a CS student.
    a. *Probe for: when was it, challenges common/other examples, severity of challenge*
15. Now I want to ask you specifically about any discrimination you have experienced. Including examples of overt bias or mistreatment, but also including intentional or unintentional subtle expressions of bias, mistreatment, or feeling uncomfortable (some people refer to these as microaggressions). Can you give me an example of a time when you experienced discrimination or a microaggression from another student or faculty?
    a. *Probe for: frequency of occurrence, ever reported an incident (easy or hard to report?). Was there any resolution?*
16. Current/Graduated student: Tell me about a time when you thought about dropping out of the program.
    a. *Probe for why the participant didn't leave, e.g., What convinced you to stay? How did you decide to stay? Who or what was instrumental in your persistence (reference prior question re support structures, mentors, etc.)?*
17. Former student: Tell me about when you first thought about leaving the program.

a. How long was it between when you first started thinking about leaving and actually leaving?
b. What was the last CS class(or classes) you took before leaving the program?
c. Did you switch to a different program of study?
d. How do you feel about this choice now?

Closing

Note: These questions will not change through the study

18. Thank you so much for your time today. This has been really valuable! Is there anything else you think I forgot to ask or that you would like to share regarding your experience with the CS program at PSU?
19. And, last, is there anyone else you would recommend I talk to for this project?

Core CS Courses

Course	At PSU or Elsewhere?	Had to retake?	Misc notes
CS 140			
CS 160 / CS 199			
CS 161			
CS 162			
CS 163			
CS 202			
CS 250			
CS 251			
CS 201			
CS 300			
CS 333			
CS 320			
CS 350			
CS 486			