

Pre-College Computing Experiences: Lessons Learned from Expansive and Inclusive Options in Surveys

Rebecca Zarch
SageFox Consulting Group
Amherst, MA
rzarch@sagefoxgroup.com

Amanda Menier
SageFox Consulting Group
Amherst, MA
amenier@sagefoxgroup.com

Megean Garvin
Maryland Center for Computing Education
at the University System of Maryland
Baltimore, MD
mgarvin@umbc.edu

Jayce Warner
University of Texas, Austin
Austin, TX
jaycewarner@utexas.edu

Bailey Brown
Spelman College
Atlanta, GA
baileybrown@spelman.edu

Phebe Palmer
SageFox Consulting Group
Amherst, MA
ppalmer@sagefoxgroup.com

Abstract—This experience report shares lessons learned when expanding demographic options on an undergraduate survey. The study is designed to better understand the relationship between pre-college computing experiences and the choice to major in computing, particularly focusing on Black women’s experiences. Expansive options for gender (5 plus an open-ended), race (18 non-mutually-exclusive options), and disability (8) gave respondents more opportunity for specificity. Yet we faced unexpected challenges in analysis and interpretation as we hadn’t considered the implications of being so expansive ahead of time. This paper presents our lessons learned, analysis choices and plans for future iterations of the survey.

Index Terms—survey, demographics, undergraduate, K–12, race and ethnicity, gender

I. INTRODUCTION

The last decade has seen rapid expansion in computer science (CS) as a core academic discipline in K–12 with widespread support from public and private agencies [1]–[3]. Unfortunately, we have learned that simply expanding CS education “for all” without purposefully attending to issues of equity across the education pipeline can sometimes exacerbate existing inequities [4] and have particularly failed Black women, who earned only 1% of computing degrees in 2018 [5]. In combating inequities, researchers and others engaged in efforts to broaden participation in computing (BPC) face two major challenges. The first is the disconnect between K–12 and postsecondary systems of education. Although they are meant to be aligned, they are functionally distinct sectors operated by completely different governing bodies and policies. As a result, very little is known about how inequities in K–12 affect participation in higher education as BPC research in each sector tends to happen independently of the other. The second major challenge, which compounds the first, is that

even within each sector tackling issues of equity is already a complex endeavor due to the myriad ways in which equity can be conceptualized and assessed. Education leaders, policy makers, and researchers have all contributed to a “cacophonous discussion of how best to define and gauge educational equity” [6].

A. Study Overview

The *Researching Early Access to Computing and Higher Education (REACH): Understanding CS pathways with a focus on Black women* research project is designed to investigate the relationship between students’ computing experiences in K–12 and higher education within a framework to assess equity across multiple sectors of education. Using statewide education data, we are conducting a longitudinal analysis that tracks students and their computing experiences across years, from 6th grade to college. This analysis will help identify which K–12 coursetaking patterns lead to more participation in computing in higher education and how these relationships differ for distinct groups of students in order to measure progress towards BPC. This analysis will be coupled with a survey of computer science majors, which is designed to understand how students experienced K–12 computing and how these experiences translate to higher education will further help refine the K–16 computing education ecosystem. Finally we will engage Black women in groups for data interpretation and personal interviews to better understand their unique experiences within the CS ecosystem.

B. Intersectionality

We center our research within the theoretical framework of Black Feminist Thought and, specifically, intersectionality. Intersectionality posits that one’s identity is not only created from race, gender, and other identifiers, but that the overlap

of those identifiers creates unique experiences that can only be understood by interrogating the intersection [7], [8]. Intersectionality also asks researchers to not solely focus on the ways individuals adapt in the face of oppression, but instead understand the systems and structures that are oppressing them [7], [8].

Applying this analytic lens to the analysis, we will utilize the domains of power framework to gain a deeper, more systemic understanding of the experiences of Black women in computer science. Using a mixed methods design, we will synthesize the quantitative and qualitative analyses to better understand the inequities that Black women experience in computing education. Because it will examine longitudinal trends across all public middle schools, high schools, and colleges in each state, the quantitative analysis will be key to illuminating where systemic inequities are manifest in coursetaking data. Because it centers the voices of Black women in computing, the qualitative analysis will be key to understanding what can be done to mitigate or eliminate these inequities.

This project is still early in its development. The first significant data collection effort has been the launch of the Pursuing Advancement in Higher Education-CS Survey (PATH-CS) to students in three institutions in Maryland: a non-flagship state university, an HBCU, and a community college.

C. Path-CS Survey

This paper focuses on the design, implementation and lessons learned from the Path-CS survey of pre-college computing experiences, particularly as it relates to the demographic questions concerning race and ethnicity. The PATH-CS survey was designed for use with computer science majors to better understand the pre-college computing experience and the connection to their decision to pursue a computing major and their experiences and sense of identity in the field. Given the purpose of the study and analytic frame, allowing students to self-describe their demographic characteristics is critical for answering our research questions.

II. DESCRIPTION OF PRACTICE: INCLUDING CONTEXT, AUDIENCE, GOALS, AND OUTCOMES

A. Design of the Instrument

The survey was substantively built upon the *Effectiveness of Technology Survey* [9] and the *CIC Baseline Computer Science Education Survey* [10], and utilized the *single item measure for assessing STEM identity* [11]. In 2020, prior to this grant, a version of the survey was piloted by two of the researchers at a university with 8,000 undergraduates to understand their K–12 experience with computing and how they may or may not be applied to their undergraduate course of study. This survey was further modified for the REACH to better align with the study goal of exploring current computer science majors' pre-college experiences. The survey was designed for use in two states, and within each state students from a non-flagship state school, a Historically Black College or University and a Community

College will be invited to participate. Ultimately, the survey included four sections:

- 1) College Experience: asks about the student's major area of study, motivation and supports
- 2) Pre-College Experience: asks about formal and informal exposure to computing prior to college and its connection to the choice to study computing
- 3) Computing Identity: asks the extent to which students see themselves as computing professionals and how they experience the computing environment
- 4) Background: asks for demographic information including social identifiers and academic descriptors.

Given the importance of understanding the relationship between identity and computing, the study team decided to be as expansive as possible with the demographic items. We felt that it was important for students to have the opportunity to reflect their identity as accurately as possible concerning gender, race and ethnicity, and disability. The team examined several other surveys for current practices in answering demographic items, including the National Academies Report *Measuring Sex, Gender Identity, and Sexual Orientation* [12], and ultimately included the following three questions in the survey:

- 1) How do you currently describe your gender identity?
- 2) What is your race/ethnicity? Please select all that apply.
- 3) Do you have any of the following disabilities or medical conditions? Please select all that apply.

The response options for each of these questions are included in tables I, II and III respectively.

In addition to the expansive and inclusive options, the demographic section of the survey included the following preamble:

A major goal of this study is to examine how people's social identifiers (gender identity, race/ethnicity and disability status) and academic status (GPA, year in school) influence their decisions to go into computing and how they experience it. We ask for certain demographic information about you in order to remain cognizant of how differences and intersections of identities can shape one's participation in computing. We thank you in advance for sharing this information with us. As a reminder, this survey is anonymous, and any reporting produced will be done on an aggregate basis.

Prior to implementation the survey was reviewed by all members of the study team and our advisory board provided feedback asynchronously.

B. Implementation

The survey was administered in Fall 2022 in the first state involved in the study. We had 450 responses to the survey with 377 responding to the gender question, 376 to the race/ethnicity question and 353 responding to the disability question (see Tables 1–3 for results)

III. LIMITATIONS AND ASSUMPTIONS

In early 2023 with a clean data set the team was eager to start with the analysis. Yet we were immediately confronted with many challenges that our desire for more expansive identity options presented.

A. Examining the data by demographics

As with most survey data, we anticipated examining our findings by subgroups. For example, were there differences in pre-college math attainment based on race? By being expansive, some categories had too few students to make any interpretation of the data. By hoping to help students feel more seen, we now risk invisibilizing them by excluding their results. For example, [refer to figure data].

We also used expansive categories to collect gender and disability data, yet we did not have the same questions about how to reduce the categories. In part, we expect that this is because we did not have a hypothesis about a specific type of disability or about gender identities that are not male or female. If we had, for example, wanted to specifically analyze our data for relationships between neurodivergence and computing identity, we would likely be having a similar conversation about which categories of disability to include in this aggregate measure, and we would likely miss some people in our sample who were neurodivergent, but did not consider it a disability, or did not disclose because they did not realize the importance of the question to the overall goals of the survey.

One option would be to collapse categories. In our race/ethnicity question, the items could reasonably be rolled up into American Indian or Alaska Native; Asian, African-American/Black, LatinX/Chicano/Hispanic; Middle Eastern; Native Hawaiian or Pacific Islander; White; Another race, ethnicity or origin. However, we cannot be confident that this will accurately represent the intentions of each student. For example, “Caribbean” will become a subcategory of “African American/Black.” Additionally, because the items were not mutually exclusive, this means that people would be represented in more than one category. We would also need to consider if we would include a category for multi-racial.

These challenges are compounded when we want to examine data by multiple demographic markers.

B. Positionality Statement

The REACH study team includes researchers with expertise in evaluation, policy, education, sociology, and includes depth in qualitative and quantitative methods. Our team is also diverse, including people with gender expansive identities, two Black women and people with disabilities. These identities, lived experiences, and diverse areas of professional expertise influenced how we conceived of the project and designed the survey instrument. Though we all reviewed the survey, and engaged our advisory board, also made of people with diverse professional expertise and lived experiences, we had not established the space and a culture within our team yet to

TABLE I
GENDER IDENTITY OF PARTICIPANTS

Response Option	Responses
Woman	118
Man	238
Two-Spirit	1
Non-binary	8
Genderqueer/gender non-conforming	4
I use a different term:	1
I prefer not to answer	7

TABLE II
RACE/ETHNIC IDENTITY OF PARTICIPANTS

Response Option	Responses
American Indian or Alaska Native	6
East Asian (e.g., Chinese, Japanese, Korean, Taiwanese)	42
Southeast Asian (e.g., Cambodian, Vietnamese, Hmong, Filipina/o/x)	24
South Asian (e.g., Indian, Pakistani, Nepalese, Sri Lankan)	58
Other Asian (Please specify)	3
African American/Black	82
African	31
Caribbean	10
Other Black (Please specify)	2
Puerto Rican	1
Central American	6
South American	7
Mexican American/Chicana/o/x	5
Other Hispanic or Latina/o/x (Please specify)	8
Middle Eastern (e.g., Arab, Middle Eastern, Persian)	4
White (e.g., European, Caucasian)	149
Another race, ethnicity, or origin not on this list (Please specify)	4
Prefer not to answer	6

TABLE III
DISABILITY STATUS OF PARTICIPANTS

Response Option	Responses
Learning disability (e.g., dyslexia, dysgraphia)	9
Attention-deficit/hyperactivity disorder (ADHD)	50
Autism spectrum disorder	11
Physical or sensory disability (e.g., speech, sight, mobility, hearing)	7
Chronic illness (e.g., cancer, diabetes, autoimmune disorders)	6
Psychological disorder (e.g., depression, anxiety, PTSD)	44
Another disability or medical condition (Please specify)	4
I have a disability but prefer not to specify the type	6
Prefer not to answer	24
Not applicable	230

move beyond the task at hand (crafting survey questions) to think about the broader implications.

Our experience is an example of why even justice-forward best practices for demographic data collection cannot be applied uncritically. Even with a team of experts and scholars, we were unable to foresee the possible limitations and questions that would arise when we reached the data analysis stage, despite what we thought was a great deal of planning and revision during survey construction. Yet, because we have diverse perspectives on our team, we are able to be reflective and cognizant of our assumptions, even if it is at a later phase, and make modifications to the next survey implementation. More importantly, our team is engaging in reexamining our assumptions when we talk about Black women.

C. Who are we including when we say “Black women?”

Our study has an explicit emphasis on understanding the experience of Black women. Our proposal justified this by citing the documented disparities faced by Black women in computing, and our analytic framework relies on the theoretical framework of Black Feminist Thought, specifically intersectionality. Yet we as a research team did not spend the time to ask ourselves “who are we including when we use the term Black women?” If we had spent the time as a team delving into this early, we may have been more cognizant of how our survey was being constructed and whether it would really give us the data we would need to a) identify the people we wish to include and b) do so in a way that allows us to then interrogate the domains of power that influence their computing trajectory.

IV. IMPLICATIONS AND NEXT STEPS

Fortunately the research team has an opportunity to learn from this experience and influence this current study. First, we as a group are going to spend time discussing our own assumptions and expectations when we use the term “Black women” in this context. We encourage other groups who are collecting demographic information to be clear on how they will use this information in their analysis. In the next iteration, there will be an additional option for people to select that they identify as a Black woman explicitly, allowing us to confidently include them in our analysis and discussion about the experiences of Black women in computing as compared to others. Our core research questions rely heavily on gender and racial identity but also assume students have been educated in the United States as we are examining K–12 through postsecondary pathways in the United States.

Second, the survey is set to launch in our second state at three new institutions in Fall 2023. We have discussed the ethics of reducing the number of groups at length internally, and we are modifying the preamble to the demographics section to include an additional explanation of how the data may be aggregated up so that students are informed of how their selection may be used and represented in our analysis and findings should we not have the statistical power necessary to examine demographic subgroups with greater specificity. No

matter the level of analysis used, how we link this information back to our framing of intersectionality and domains of power will continue to be an important conversation for the team. For example, the demographic supercategory of “Asian” groups together students from a multitude of ethnic and religious backgrounds who may have very different experiences in home life or institutional support [13].

Finally, the survey is only one part of the study. By engaging our diverse team in interrogating our own assumptions and their impact on our design choices, our future efforts will benefit. Future stages of the project will bring Black women in computing majors to help us interpret findings. These findings will also be augmented with in-depth interviews and focus groups with Black women in computing. The lessons learned in this phase of the study point to the benefit of qualitatively assessing students’ concepts of identity and background as part of their computing experience during these sessions.

V. ACKNOWLEDGMENT

This material is based upon work supported by the U.S. National Science Foundation under Grant No. 2201700. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We’d also like to thank study team members Dr. Tamara Pearson for helping frame the study and Talia Goldwasser for producing this piece.

REFERENCES

- [1] J. Cuny, “Transforming K–12 computing education: an update and a call to action,” *ACM Inroads*, vol. 6, no. 3, pp. 54–57, 2015, doi: 10.1145/2809795.
- [2] Google LLC and Gallup Inc. “Searching for computer science: Access and barriers in U.S. K–12 Education.” 2015. URL: https://services.google.com/fh/files/misc/searching-for-computer-science_report.pdf.
- [3] H. Partovi, “A comprehensive effort to expand access and diversity in computer science,” *ACM Inroads*, vol. 6, no. 3, pp. 67–72, 2015, doi: 10.1145/2807704.
- [4] J. Margolis, R. Estrella, J. Goode, J. Jellison Holme, and K. Nao, *Stuck in the Shallow End: Education, Race, and Computing*. Cambridge, MA: MIT Press, 2008.
- [5] Y. Rankin and J. O. Thomas, “The intersectional experiences of black women in computing,” in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 2020, pp. 199–205.
- [6] National Academies of Sciences, Engineering, and Medicine, *Monitoring Educational Equity*. Washington, DC: The National Academies Press, 2019, p. 35, doi: 10.17226/25389.
- [7] P. H. Collins and S. Bilge, *Intersectionality*, 2nd ed., Cambridge, UK: Polity Press, 2020.
- [8] K. Crenshaw, “Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics,” *University of Chicago Legal Forum*, vol. 1989, 1989, Art. no. 8. URL: <http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>.
- [9] M. M. McGill, A. Decker and A. Settle, “Undergraduate students’ perceptions of the impact of pre-college computing activities on choices of major,” *ACM TOCE*, vol. 16, no. 4, 2016, Art. no. 15.
- [10] Momentum, “The CIC baseline computer science education survey,” University of California Los Angeles, 2021, unpublished.
- [11] M. M. McDonald, V. Zeigler-Hill, J. K. Vrabell, and M. Escobar, “A single-item measure for assessing STEM identity,” *Frontiers in Education*, vol. 4, 2019, p. 78, doi: 10.3389/educ.2019.00078.

- [12] National Academies of Sciences, Engineering, and Medicine, *Measuring sex, gender identity, and sexual orientation*. Washington, DC: The National Academies Press, 2022, doi: 10.17226/26424.
- [13] S. Toppa and D. Chambers. "Students labeled 'Asian' are not a monolith. Why are schools treating them that way?" TFA One Day. <https://www.teachforamerica.org/one-day/policy-and-advocacy/students-labeled-asian-are-not-a-monolith-why-are-schools-treating-them> (accessed May 15, 2023).