

Examining the Computing Identity of High-Achieving Underserved Computing Students on the Basis of Gender, Field, and Year in School

Authors: Atalie Garcia, FIU. Monique Ross, FIU. Zahra Hazari, FIU. Mark Weiss, FIU. Kenneth Christensen, USF. Michael Georgiopoulos, UCF.

ABSTRACT

As technology increases in the global arena and the necessity for a more diverse group of individuals to fulfill engineering and computing roles increases, it is important to engage more students in computing majors and roles. Identity has proven to be an important lens through which researchers can better understand how to engage students in these fields. In particular, our framing for computing identity includes students' self-perceptions about recognition, interest, and performance/competence. Using survey data, this study examines the computing identity of high achieving underserved students in computer science (CS), computer engineering (CE), and information technology (IT). For these students, we compare the constructs between men and women, computing fields, and freshmen and juniors/seniors. Based on preliminary data, results show that female participants had less of a computing identity than male students, specifically with respect to computing recognition and overall computing identity. Students in IT programs had less of an overall computing identity than students in CS and CE. Finally, freshmen were lower on their overall computing identities and specifically performance/competence beliefs and interest. These results suggest that even within computing programs, students differ in their computing identities. Furthermore, there are different constructs related to computing identity that vary for different subgroups of students. While these results are insightful, future work will compare the computing identities of high achieving underserved students with a larger population of computing students as well as strategies for building their computing identities.

1 INTRODUCTION

There has been a growing movement advocating that computer science should be a core competency for students in the US. For example, the President's office in 2016 issued a statement about "Computer Science for All" to engage all students in CS so that they are equipped with the computational thinking and skills needed to be innovators in our technology-driven society [1]. The Department of Labor also estimates a large growth in job demand for careers related to high technology sectors [2]. Despite this fact, there continue to be students who are marginalized in CS and lack the resources and support to have the same opportunities for access and success within these fields. For example, women, underrepresented race/ethnic groups, and economically underprivileged groups have all been found to be disadvantaged in computing fields in ways such as mathematics preparation, technology access, role models, and attitudes/stereotypes [3]. This paper sets out to better understand the attitudes of these types of "underserved" students' who are in computing disciplines but who are also identified as high achieving. While much work compares underserved students to more typical normative students [4], we examine attitudes within a group of underserved students who are successful at their stage within computing programs. As such, this work brings a novel perspective to the research in computer science education. In particular, we examine how students see themselves with respect to computing (computing identities) and how these perceptions vary within this group. We approached this research study with identity as the guiding framework to answer the following research questions:

1. How do the computing identities of high achieving underserved students differ by gender?
2. How do their computing identities differ by field (CS, IT, CE)?
3. How do their computing identities differ by freshmen versus upper class status?

2 BACKGROUND

National attention to retaining U.S. prominence on the global stage has precipitated a necessity to garner interest, enrollment, and subsequent graduation of computing majors. The demand for a technologically trained workforce far outpaces the growth. After all, computing is projected to continue to grow at a rate of 17% from 2014-2024, much faster than the average for all occupations [2]. This urgency has prompted an expansion in literature on understanding the engagement of women, underrepresented minorities, and diverse socio-economic backgrounds in computing fields.

Research surrounding women's engagement in computing has been on the rise in recent years. It began with the realization that computer science is the only STEM field that is experiencing a steady decline of female enrollment since the 1980's, 37% to 18% [5]. Since this revelation, rigorous research has highlighted the barriers to computing which include environment and climate, stereotypes, and self-efficacy, to name a few [6-8]. Exploration has also included initiatives by various organizations and universities that have proven to be successful at attracting and retaining women in computer science [9-10].

Another demographic with paltry representation within the fields of computing garnering attention in the scholarly community are underrepresented race/ethnicity groups (URG) (including sub-populations such as black and Hispanic). URG have consistently engaged in computing at lower rates with Hispanics earning 6.7% of the computer science bachelor's degrees and African-American/Blacks earning 5% [11]. Scholars have noted factors such as lack of role models, interest, and access as primary factors for low engagement [3]. This topic of access often refers to the digital divide or the huge disparity between those that have access to computers and computer literacy and those that do not. Despite the decreasing cost of the personal computer the financial commitment is still a necessity and this financial burden makes acquiring computers a luxury of the households/families that can afford them. The same financial challenge is prevalent in schools and educational institutions. This limits access to computers and similar technologies for those students from lower socio-economic status (SES) [12-13]. Computing fields, however, are not unique in their need to reach a broader audience and engage a more diverse population. Other STEM fields, including engineering, have also begun to evaluate and explore opportunities to engage more women and URG into their disciplines. Research in this space has established relationships between the construct of identity and engagement and persistence [14-17].

3 THEORETICAL FRAMEWORK

Identity is a theoretical lens that has been utilized effectively in science and engineering education to better understand engagement and persistence of STEM students. Identity has been measured by the ways that students talked about STEM, acted or participated in class, described themselves within the context of the STEM world (or classroom), and related to others in the

STEM community. In this framing, disciplinary identity is described as the ways in which students envision themselves in the context of a domain or discipline [18]. Incongruence between how they describe a STEM community and how they envision themselves within that space can contribute to their disengagement from the STEM learning community (i.e., within classrooms and programs). The construct of identity has been demonstrated by scholars [19] as including the sub-constructs of recognition, interest, and performance/competence. For this study, the constructs of performance and competence have been combined following the work of science identity scholars that determined that while experts in STEM can distinguish the difference between performance and competence, undergraduate students cannot [22]. A student's disciplinary identity as measured by examining their interest, perceived performance/competence, and whether they believe they are being recognized by others as being a part of the discipline has been found to be strongly predictive of their persistence in that discipline [17]. This theoretical framing has recently been validated for examining computing identity in similar ways to STEM, i.e., the computing identity measure built from the constructs of recognition, interest, and performance/competence is strongly predictive of the choice of a computing career [20]. These findings are still under review and therefore mentioned as recognition of work in progress in this space. However, all references to the role of identity in persistence will largely be from the body work established in STEM identity.

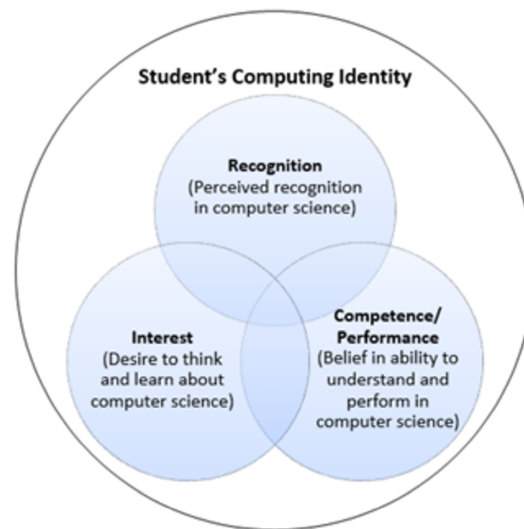


Figure 1. Computing Identity Framework

4 DATA AND METHODS

The data was collected as part of a large National Science Foundation (NSF) funded project, [title and grant number blinded for anonymity], geared towards enhancing the educational experiences of high-achieving underserved students in Florida. The project focuses on supporting freshmen and rising seniors. Data was collected from full-time high-achieving underserved students at three participating institutions enrolled in degree programs in computer engineering (CE), computer science (CS), and information technology (IT). High-achieving underserved students were defined as those who qualified for Federal Student Aid and who also maintained a required minimum GPA (3.0 for freshmen, 2.5 for upcoming seniors). A GPA of

2.5 was considered to be high achieving for senior students due to the threshold proposed by the S-STEM grant for scholarship eligibility.

The survey instrument was designed and developed by leveraging items from valid and reliable instruments in science and computing education [19, 20] as well as performing additional reliability and validity testing. Feedback was also received from education researchers and measurement experts for content validity. In terms of construct validity for the computing identity measures, exploratory factor analysis (EFA) with promax rotation (the theorized sub-constructs are oblique, i.e., overlap) was used to assess the structural validity of the computing identity measures (interest, recognition, performance/competence).

An exploratory factor analysis (EFA) revealed that the survey included four items related to recognition, four items related to interest, and two items related to performance/competence beliefs that loaded onto each factor with a loading of 0.4 or greater. The cumulative variance explained by the three factors in the EFA was 66.1%. Identity sub-construct sample questions are listed in Table 1. Items in each factor were averaged to create overall measures for computing recognition, computing interest, and computing performance/competence beliefs. These three factors were then averaged for an overall computing identity measure.

Upon institutional review board (IRB) approval, preliminary data was collected spring 2017 on their computing identity as conceptualized by the theoretical framework. In total, 95 students were solicited and all of them responded to the survey. However, background information (e.g., gender) was available for only 93 (two respondents did not provide correct IDs so their background information could not be retrieved). The complete data from the 93 respondents was analyzed for this paper. Of the sample, 12 were black non-Hispanic, 27 white non-Hispanic, 39 Hispanic, and 15 who identified otherwise. Note, due to a limited sample size amongst black non-Hispanic students, we chose not to make group comparisons by race/ethnicity. This will be done in a subsequent study where the sample sizes are larger. Furthermore, our data included responses from: 23 women and 70 men; 17 CE, 47 CS, and 29 IT students; 44 freshmen and 49 upcoming/current seniors. These latter groups were compared using Kruskal-Wallis non-parametric tests (due to uneven sample sizes) on the computing identity measures described below.

Table 1. Identity sub-construct sample questions

Recognition: Please rate the following statements as they apply to you.

I see myself as a computer savvy person

My family sees me as a computer savvy person

My friends/classmates see me as a computer savvy person

My instructors/teachers see me as a computer savvy person

Interest: Please rate the following statements as they apply to you.

Topics in computing excite my curiosity

Computer programming is interesting to me

I enjoy learning about computing

I like to know what is going on in computing

Performance/competence: Please rate the following statements as they apply to you.
 I can do well on computing tasks (e.g., programming and setting up servers)
 I understand concepts underlying computer processes

Note. Response options for each item were a 5-point anchored scale: 0 = Not at all; 4 = Very much so

5 RESULTS

The overall computing identity measure for the group of high-achieving underserved students surveyed had a mean value of 3.3 (± 0.6 SD). In terms of the identity factors, the mean for recognition was 3.3 (± 0.6 SD), interest was 3.5 (± 0.6 SD), and performance/competence was 3.0 (± 0.8 SD). Table 2 summarizes the means and standard deviations for female and male students and indicates whether these means are significantly different using a Kruskal Wallis test. When comparing female and male students, female students were significantly lower in their sense of recognition and overall computing identity. Although the means for interest and performance/competence are also lower for female students, these differences were not significant at the $p < 0.05$ level.

Table 2. Computing Identity Means (with Standard Deviations) by Gender

	All Students	Female	Male	Sig. Diff. [†]
Recognition	3.3 (0.6)	3.0 (0.6)	3.4 (0.6)	0.005 (**)
Interest	3.5 (0.6)	3.2 (0.7)	3.6 (0.5)	0.11 (ns)
Perf./Comp.	3.0 (0.8)	2.8 (0.6)	3.0 (0.8)	0.08 (+)
Computing Identity	3.3 (0.6)	3.0 (0.6)	3.3 (0.5)	0.02 (*)

[†] p-values for Kruskal Wallis Test

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Next, we compared computing identity by field, specifically Computer Engineering (CE), Computer Science (CS), and Information Technology (IT). Table 3 summarizes the means and standard deviations by field and the results of the Kruskal Wallis tests. The only significant result is for the overall computing identity measure where IT students were found to be significantly lower than CS students.

Table 3. Computing Identity Means (with Standard Deviations) by Field

	CE	CS	IT	Sig. Diff. [†]
Recognition	3.4 (0.5)	3.4 (0.6)	3.1 (0.7)	0.12 (ns)
Interest	3.6 (0.5)	3.6 (0.6)	3.3 (0.6)	0.06 (+)
Perf./Comp.	2.8 (0.8)	3.1 (0.8)	2.8 (0.8)	0.17 (ns)
Computing Identity	3.3 (0.4)	3.4 (0.6)	3.1 (0.6)	0.04 (*)

[†] p-values for Kruskal Wallis Test

+ $p < 0.10$, * $p < 0.05$

Lastly, we compared computing identity by year, specifically first year students and rising seniors. Table 4 summarizes the means and standard deviations for these groups and the results of the Kruskal Wallis tests. The results show that first year students are significantly lower in their interest, performance/competence beliefs, and overall computing identity. Note that across all of the analyses by group, the lowest mean value is for first year students in their computing performance/competence beliefs.

Table 4. Computing Identity Means (with Standard Deviations) by Year

	1Y (Freshman)	3YearPlus (Juniors/Seniors)	Sig. Diff. [†]
Recognition	3.2(0.7)	3.4 (0.6)	0.09 (+)
Interest	3.4 (0.6)	3.6 (0.5)	0.03 (*)
Perf./Comp.	2.6 (0.8)	3.3 (0.7)	0.0001 (***)
Computing Identity	3.1 (0.6)	3.4 (0.5)	0.002 (**)

[†] p-values for Kruskal Wallis Test

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

6 DISCUSSION AND CONCLUSION

Although there is an extensive amount of literature written about the disparity between gender, race, ethnicity, and age in the computing profession [4,18,21], it serves the computing community to look at more niche groups when analyzing computing identity. In this study, we looked at students who are high-achieving and underserved in computing. However, not all these students persist. Looking at high-achieving and underserved students presents us with the ability to see, specifically, where students of different genders, discipline, and year in school differ in their computing identity.

The first research question guiding this study, explored the differences in gender when examining computer science identity. There is a clear difference between genders when it comes to computing identity, specifically in recognition (males scored 3.4 overall while women scored 3.0). This shows that women who are high achieving in computing still show signs of feeling less acknowledged as computing people than male students. This means that, at home, at school, and in social circles, women do not feel as if they are being recognized as having knowledge in computing. Recognition is a strong indicator within the STEM identity construct of a future career in STEM [19]. It is concerning that the women in our study that have been identified as being high achieving, feel less recognized due to the fact that it may impede on their rates of persisting to a career in computing. It would prove beneficial to further study the factors involved with lowered recognition for women and possibly see what ways institutions and communities can change in order to improve this perception.

The second research question guiding the study, explored the difference between computing disciplines when examining computing identity. There were notable contrasts with regards to interest across the disciplines. Interest is defined as a student's desire to learn and know computing and follows recognition in having strong predictive power for a student's likelihood of pursuing a STEM career [19]. Curiously, IT scored lower overall in interest compared to its

counterparts, with a difference of 0.06. Although it is not statistically significant with this sample size, this may indicate that IT students do not feel as interested in their major as computer science and computer engineering students. Further exploration, with a larger population, might yield a statistical significance. However, given that IT and CS both study software, it is interesting that there is a drop of interest between these majors. This raises the following questions: what influences this lack of interest? Could it be an issue of perception? As previously mentioned, the high demand in computing fields makes it prudent to understand why there is a difference in interest and identity between information sciences and computer science.

The third and final research question explored computing identity between students in the various stages of degree completion. It should be noted that computing identity and performance/competence had a significant difference when it came to school year. In looking at the trends, as students matriculated in their degree programs, it was understandable that interest, recognition, and overall identity, increased. The result that yielded more questions than answers was the construct of performance and competence as it pertains to freshman versus upper level classmates. The first-year student average of 2.6 versus the far more confident 3.3 of upper level classmates could indicate that curriculum is designed to build perceived competence or that the curriculum has weeded out those with lower competence. The statistical significance suggests that performance and competence should be further examined along the trajectory (freshman, sophomore, junior, and senior) to better understand student perceptions of themselves with regards to competence in the area of computing. The research team initially assumed that students identified as high achieving would have higher performance/competence in the area of computing; however, this was not the case.

We often focus on underprepared students and their self-perceptions in computing fields, however this study unveils that even students identified as high achieving are at risk of possessing a lack of perceived competence and performance. Future work will include tracking these high achieving, low perceived performance and competence students, to better understand the implications of the sub-construct on their prolonged engagement.

Further study with a larger sample size would be beneficial for the computing community to further analyze these constructs on a larger scale to better understand the implications that these constructs have on students of different genders, underrepresented race/ethnic groups, and economic backgrounds. A qualitative exploration into factors that foster and strengthen computing identity salience beyond the survey items would also provide complementary and rich insight into the retention and/or attrition of computing students.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1643965. 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.

REFERENCES

- [1] White House., 2016. Computer science for all., Retrieved from <https://www.whitehouse.gov/blog/2016/01/30/computer-science-all>.
- [2] Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, 2016-17 Edition, Computer and Information Research Scientists, Retrieved from <http://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-Scientists.htm>
- [3] Buzzetto-More, N., Ukoha, O. and Rustagi, N., 2010. Unlocking the barriers to women and minorities in computer science and information systems studies: Results from a multi-methodological study conducted at two minority serving institutions. *Journal of Information Technology Education: Research*, 9(1), pp.115-131.
- [4] Scherr, R., 2016. Never mind the gap: Gender-related research in Physical Review Physics Education Research, 2005–2016. *Physical Review Physics Education Research*, 12(2), p.020003.
- [5] Accenture., 2016. Cracking the gender code: Get 3x more women in computing. Retrieved from <https://www.accenture.com/us-en/cracking-the-gender-code?src=JB-11540>
- [6] Cohoon, J.M., 2001. Toward improving female retention in the computer science major. *Communications of the ACM*, 44(5), pp.108-114.
- [7] Cheryan, S., Plaut, V.C., Davies, P.G. and Steele, C.M., 2009. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology*, 97(6), p.1045.
- [8] Vekiri, I. and Chronaki, A., 2008. Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & education*, 51(3), pp.1392-1404.
- [9] Frieze, C. and Quesenberry, J., 2015. *Kicking Butt in Computer Science: Women in Computing at Carnegie Mellon University*. Dog Ear Publishing.
- [10] NCWIT., 2010. Harvey Mudd College's successful systemic approach. Retrieved from <https://www.ncwit.org/resources/how-does-engaging-curriculum-attract-students-computing/harvey-mudd-college%E2%80%99s-successful>
- [11] Bock, S. J., Taylor, L., Phillips, Z., & Sun, W., 2013. Women and minorities in computer science majors: Results on barriers from interviews and a survey. *Issues in Information Systems*, 14(1), 143–152.
- [12] Bozionelos, N., 2004. Socio-economic background and computer use: the role of computer anxiety and computer experience in their relationship. *International Journal of Human-Computer Studies*, 61(5), pp.725-746.
- [13] Ritzhaupt, A.D., Liu, F., Dawson, K. and Barron, A.E., 2013. Differences in student information and communication technology literacy based on socio-economic status, ethnicity, and gender: Evidence of a digital divide in Florida schools. *Journal of Research on Technology in Education*, 45(4), pp.291-307.
- [14] Tonso, K.L., 2006. Student engineers and engineer identity: Campus engineer identities as figured world. *Cultural studies of science education*, 1(2), pp.273-307.
- [15] Tate, E.D. and Linn, M.C., 2005. How does identity shape the experiences of women of color engineering students?. *Journal of Science Education and Technology*, 14(5), pp.483-493.

- [16] Hughes, R.M., 2015. An investigation into the longitudinal identity trajectories of women in science, technology, engineering, and mathematics. *Journal of Women and Minorities in Science and Engineering*, 21(3).
- [17] Godwin, A., Potvin, G., Hazari, Z. and Lock, R., 2016. Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice. *Journal of Engineering Education*, 105(2), pp.312-340.
- [18] Kane, J.M., 2012. Young African American children constructing academic and disciplinary identities in an urban science classroom. *Science Education*, 96(3), pp.457-487.
- [19] Hazari, Z., Sonnert, G., Sadler, P.M. and Shanahan, M.C., 2010. Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47(8), pp.978-1003.
- [20] Mahadeo, J., Hazari, Z., Sonnert, G., & Sadler, P.M. 2017. (In revision). Developing a Computing Identity Framework. Manuscript in revision for *IEEE Transactions on Education*.
- [21] Master, A., Cheryan, S. and Meltzoff, A.N., 2016. Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology*, 108(3), p.424.
- [22] Potvin, G., & Hazari, Z. (2013). The development and measurement of identity across the physical sciences. *Physics Education Research Conference (PERC) Proceedings*, Portland, OR.