Developing and Validating a Scale on Subjective Factors Affecting Retention of Underrepresented Computer Science Undergrads

Oluwafemi Osho
Clemson Univesity
Clemson, SC, U.S.A
oosho@clemson.edu

Ananya Gupta
Clemson Univesity
Clemson, SC, U.S.A
ananyag@clemson.edu

Devon Brown
Howard University)
Washington, DC, U.S.A
devon.brown@bison.howard.edu

Bart Knijnenburg
Clemson Univesity
Clemson, SC, U.S.A
bartk@clemson.edu

Cazembe Kennedy
Vanderbilt University
Nashville, TN, U.S.A
cazembe.kennedy@vanderbilt.edu

Gloria Washington
Howard University
Washington, DC, U.S.A
gloria.washington@howard.edu

Stacey Sexton
Sagefox Consulting Group
Amherst, CA, U.S.A
ssexton@sagefoxgroup.com

Lingyuan Li
Clemson Univesity
Clemson, SC, U.S.A
lingyu2@clemson.edu

Eileen Kraemer
Clemson Univesity
Clemson, SC, U.S.A
etkraem@clemson.edu

John Porter III
Morehouse College
Atlanta, GA, U.S.A
John.Porter@morehouse.edu

Kinnis Gosha
Morehouse College
Atlanta, GA, U.S.A
kinnis.gosha@morehouse.edu

2024 Black Issues in Computing Education (BICE) | 979-8-3503-7291-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/BICE60192.2024.00010

***Abstract*—This paper discusses research aimed at developing, validating, and improving the robustness of survey scales to assess the subjective factors that impact the retention and graduation rates of underrepresented minorities (URMs) in computer science (CS) undergraduate programs. Our iterative scale development process consisted of: 1. Selecting questions from existing CS survey instruments; 2. Administering the survey at three different institutions; 3. Conducting exploratory and confirmatory factor analyses on the survey items; and 4. Updating the items based on the analysis results. Although a few of the scales are still a work in progress, after two iterative rounds with three surveys (N = 184, N = 338, N = 450), we established fourteen robust factors that can be adopted in future studies. The factors measure students' attitude toward CS, intention to continue their CS trajectory, three factors regarding students' beliefs about computer scientists, computing identity, current experience, proactive and preemptive help-seeking, familiarity with future opportunities, persistence, leadership, confidence, and perception of social support. Further, we demonstrate that these scales can be used to uncover differences between different groups of students (e.g., men and women, Black and non-Black students).**

***Index Terms*—Societal factors, Underrepresented minorities, URMs, diversity, scale development**

I. Introduction

This paper addresses the still limited understanding regarding the factors affecting the representation and retention of women and underrepresented minorities (URMs), such as Black students in Computer Science (CS). Representation of such students is found to be low [1]–[10], and their dropout rates remain elevated [3]–[5], [11]. Existing research indicates that various factors contribute to the higher dropout rates and lower representation of URMs in CS programs, including academic aspects such as course difficulty and teaching quality [12]–[14], social and community elements such as inclusivity and mentorship [1], personal factors such as interest, passion, self-efficacy, and time management [15], [16], as well as broader societal and industry trends like relevance and job opportunities [4]. Among these factors, subjective factors have been shown to significantly impact the experiences and retention of URMs in CS [4]. For instance, previous studies have highlighted subjective issues such as the lack of inclusivity, particularly for African American women [1], and the perceived relevance and utility of CS degrees [4].

While past research identifies some subjective factors, such as self-efficacy, interest, students' prior experience in programming, and sense of belongingness as well-studied factors [11], [16], [17], that may influence the retention of women and URMs in CS, there may be additional factors that remain to be uncovered, especially given the rapidly evolving nature of CS studies and societal expectations. Consequently, a comprehensive exploration of potential subjective factors that can promote URM retention and success in CS along with investigating influences of such factors on URM students with a large sample size [16], particularly at the undergraduate level, is imperative. To that end, our research involves a collaboration between two Historically Black Colleges and Universities (HBCUs), Morehouse College and Howard University, and a Primarily White Institution (PWI), Clemson University. These HBCUs provide access to a significant population of undergraduate students from underrepresented

This research was supported by NSF award DUE 2111354.

minorities (URMs). Additionally, we present an iterative procedure in which we build upon existing research, develop our own scales, and empirically validate them through three survey studies (N = 184, N = 338, N = 450). These studies produce a set of robust scales that can help academic institutions and policymakers understand the core elements influencing URM undergraduates' retention and graduation rates. Furthermore, these scales can contribute to creating an enhanced academic and socially supportive environment, fostering the retention of undergraduates from diverse backgrounds in the field of CS.

II. Related Works

A. Factors Affecting Retention of URMs in CS

The lack of URM student representation in the field of CS continues to plague graduation and retention for students in CS. A 2019 national study [18] found that of undergraduates in CS programs, an estimated 19% were Women, 23% Asian, 5% Black, 11% Latinx, 45% White, and the remaining comprised of multiracial and international youth, revealing disparities in representation. When looking at the Black population in the United States in 2019, Black people were 12% of America. This stark gap between Black students in CS degrees and society emphasizes the lack of representation.

Subjective Factors and Influences: Research has delved into societal factors that might contribute to these disparities. Research from 2010 [19] highlights societal factors that could influence these low enrollment rates, such as the lack of role models and exposure to family members succeeding in technology-related jobs [17]. Additionally, studies show higher retention and graduation rates among students with access to social and academic integration [4], [17].

Barriers to URM Enrollment and Success: Surveying students from two HBCUs, Buzzetto-More et al. [19] identified that inadequate guidance for CS and information systems (IS) students, lack of experience studying computing prior to college, and lack of exposure to programming were barriers for URM enrollment, success and retention in the CS field. Biggers et al. [17] found that students switch to non-CS majors due to poor teaching and advising, harsh grading, and heavy demands. Another study found that sense of belonging increases with social interaction with other CS students for both male/female students and for those who do and do not identify as minorities [17], [20]. This suggests that students who may not have had the opportunity to communicate, work, or socialize with CS or technology-focused students may not feel connected to the community [11], [17].

As these persistent challenges to URM communities continue in computer science programs, there is a need to understand how these societal factors affect recruitment, sense of belonging, retention, and graduation rates.

B. Scale Development

One existing scale development effort concerns the Computer-Email-Web (CEW) Fluency scale to assess students' general understanding of computer concepts when using the World Wide Web. Bunz [15] refined the CEW fluency scale by iterating over multiple pilots. During each pilot, they employed factor analysis, sub-scales, and correlation/regression analysis to validate their processes in the CEW fluency scale [15]. Developed in 2001, this scale was rooted in familiarity with concepts that today seem rudimentary. Still, fine-tuning the scale and determining comprehension remains pertinent.

A 2021 study delved into scale development and validation concerning competence-related self-concepts to discern the predictors of performance, behavior, motivation, and well-being [21]. To achieve this, a 25-item scale was created to measure general and domain-specific information and communication technology (ICT) skills. One notable gap in the study was that even though the researchers based their model on a nomological network of technological constructs, they did not examine the relation between those constructs against self-concept scales.

Our research contributes to this evolving body of knowledge by synthesizing a collection of scales that measure subjective factors related to retention in CS education, particularly for URM CS students.

III. Study 1

A. Method

To assess the various societal factors that impact the retention and graduation rates for underrepresented minorities (URMs), we conducted a survey in the spring of 2022. To develop a subjective factors questionnaire, we examined a multitude of existing surveys in the CS Education Research survey repositories[1], filtering for the computing domain, undergraduate population, and validity and reliability. We selected a subset that would be applicable and relevant to the constructs we wanted to measure. For questions that were developed for a non-CS context, we modified them to reflect CS concepts and terminologies. We collected data from 184 undergraduate computer science students, comprising 162 students from an R1 Primarily White Institution (PWI) and 22 from an HBCU.

B. Results

1) Measurement Validity: A confirmatory factor analysis (CFA) [22] of the survey responses failed to replicate many of the existing scales. The team discussed the survey and recategorized some items using the card sorting technique [23]. We then performed a series of exploratory factor analyses (EFAs) using Oblimin rotation [24] of the problematic items to explore their dimensionality: (1) The EFA of the *Belief about computer scientists* did not produce robust results, and neither did the factors for current experience, prior experience, and students' behavioral intention to continue their CS trajectory. (2) In the EFA of the *Attitude toward CS* and *Computing identity* factors (Table I), the former construct split into two factors: *Lack of relevance of CS* and *Attitude towards CS*. (3) In the EFA of *Effort*, *Grit*, *Self-efficacy*, and *Social support* factors, the structure remained, but several items were removed due to low loading or cross-loading. Note that factors for familiarity

[1]https://csedresearch.org/

TABLE I: *Attitude towards CS* and *Comp. identity* EFA Results

	F1	F2	F3
My goals do not require that I learn computer science skills		**0.574**	
Programming is of no relevance to my life		**0.725**	
I see computer science as a subject I will rarely use in my daily life		**0.743**	
I hope that in the future, I can find a career that does not require the use of computer science	-0.177	**0.619**	
It would make me happy to be recognized as an excellent student in computer science	0.138	-0.134	**0.605**
Being regarded as smart in computer science would be a great thing			**1.015**
In general, being a computer scientist is an important part of my self-image	**0.611**	-0.121	
I have a strong sense of belonging to the community of computer scientists	**0.888**		
I have come to think of myself as a "computer scientist"	**0.784**		
Overall, I feel like I belong in computer science	**0.631**	-0.147	0.121

with future opportunities, persistence, and leadership already produced acceptable results in the original CFA, so these factors were not subjected to an EFA.

Finally, we conducted another CFA to confirm the nine factors uncovered in the EFAs, measuring students' perceived lack of relevance of CS, attitude towards CS, computing identity, study effort, familiarity with future opportunities, persistence, leadership, confidence, and perceptions of social support. This CFA demonstrated a reasonable[2] fit: $\chi^2(491)$ = 794.978, $p < .001$, CFI = 0.956, TLI = 0.950, RMSEA = 0.059 (with 90% confidence interval: [.052, .067]). Moreover, all of the factors' Average Variance Extracted values (AVEs) exceeded the recommended threshold of 0.50, demonstrating satisfactory convergent validity. To confirm discriminant validity, we ensured that the square root of the AVE for each factor surpassed its correlations with other factors.

2) Differences Between Groups: We found that women displayed lower levels of **computing identity** (9.53 v. 10.91 on a scale of 4 to 16), $t(76.7) = 2.73$, $p = .008$, $d = 0.622$) compared to men. This discrepancy might account for their lower **confidence** levels (24.20 v. 25.92 on a scale of 8 to 32), $t(80.8) = 2.88$, $p = .005$, $d = 0.641$. We also observed that Black students showed greater **familiarity with future opportunities** (3.73 v. 3.02 on a scale of 4 to 8), $t(42.9) = 2.15$, $p = .037$, $d = 0.657$) than non-Black students. Intriguingly, the HBCU students had a higher level of familiarity with future opportunities than Black PWI students (4.00 v. 3.31, ns).

IV. Study 2

A. Method

Study 2 was conducted to ascertain the robustness of the factors that worked well in Study 1 and reassess those that needed improvement. Data was collected in Fall 2022 from undergraduates taking CS courses in the R1 PWI surveyed in Study 1.

A total of 349 students participated in the survey. We included an attention-check question to ensure the participants paid attention and read the questions carefully. The data of 11 students were dropped for failing that question. Consequently, the data of 338 participants were used for analysis.

Most students identified as men (76.04%) and were non-Black (92.90%). More than half of the participants (58.28%) were enrolled in a CS1 course, while the remaining participants were in a CS2 course.

1) Subjective Factors Scale: Generally, we retained the items from Study 1 for factors that performed well if the number of items with high loadings was more than 3. For those that did not meet this requirement, we expanded the items. Hence, we added 3 additional questions from [26] for the *Leadership* factor and adapted 4 questions to the *Social support* factor from [27], [28].

In study 1, the *Attitude towards CS* constructs split during EFA into two factors: (*Lack of relevance of CS* and *Attitude towards CS*). However, the team decided to combine both constructs under the *Attitude towards CS* to test the robustness of the scales under a larger sample size.

In the case of the study effort factor, only 2 of the 4 items loaded well. We observed that the two questions bordered on seeking and requesting help. Consequently, we renamed the factor as *Asking for/receiving help*. Three new questions were created and added to the factor. We also added help-related questions that had been dropped from the *Social support* factor due to their poor performance. We surmised they might load well under the help factor.

For the *Current experience* factor (one of the four that lacked construct validity), the team felt the two constructs (used in Study 1) were a useful benchmark, so we retained them with the intent of converting them into a composite score if a factor would not be appropriate. For *Prior experience*, since none of the items worked, it was decided that the factor should be dropped.

The *Belief about computer scientists* factor was initially composed of semantic differential questions [29]. Since these were the only questions with such a structure, the team thought the students did not cope well with them, hence the weak loadings. As a result, one end of the bipolar pair was used for each question in the new survey, balancing the positively and negatively worded question parts.

The last factor with a poor scale was *Behavioral intention*. We believed the items were of poor quality. For instance, the first question presented participants with two response options: Yes or No. On the other hand, the last two questions had the following options: N/A, very unlikely, somewhat unlikely, somewhat likely, and very likely. In Study 2, we rephrased the questions and used a 5-point Likert-type response ranging from unlikely to likely. We also formulated three additional questions. The new items are asterisked in Table III.

In summary, the scale consisted of the following latent variables (with the corresponding number of questions): Attitude

[2]A significant χ^2 value shows that the model shows a significant deviation from perfect fit, which is common for factor models. Common thresholds for alternative fit statistics are CFI and TLI ≥ 0.90, $RMSEA \leq 0.08$ with the 90% CI not exceeding 0.10 [25], and $\chi^2/df \leq 3$ [4].

TABLE II: Reliability and validity table. The blue-colored cells relate to Study 3.

	CA	AVE	CA	AVE	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Attitude toward CS	0.92	0.657	0.92	0.662		0.868	0.009	-0.076	0.519	0.777	-0.447		-0.37	0.211	0.237	0.07	0.717	0.141
2. Behavioral intention	0.857	0.745	0.872	0.808	0.926		-0.113	0.106	0.463	0.787	-0.36		-0.278	0.224	0.17	0.086	0.709	0.095
3. Belief in the struggle of computer scientists	0.627	0.467						-0.644	-0.148	-0.161	0.267		0.301	-0.127	0.059	0.191	-0.269	0.086
4. False belief in the effortless process of computer scientists	0.752	0.694	0.712	0.652	-0.066	-0.053			0.063	0.208	-0.138		-0.076	0.101	-0.017	-0.037	0.15	0.057
5. Belief in the insight of computer scientists	0.745	0.667	0.762	0.686	0.332	0.262		0.108		0.462	-0.321		-0.200	0.238	0.352	0.204	0.654	0.261
6. Computing identity	0.895	0.768	0.886	0.749	0.754	0.745		0.063	0.219		-0.342		-0.143	0.337	0.287	0.14	0.689	0.252
7. Current experience	0.755	0.737	0.71	0.68	-0.301	-0.266		-0.103	-0.301	-0.23			0.464	-0.134	0.065	0.18	-0.602	-0.004
8. Proactive help-seeking			0.598	0.48	0.182	0.145		-0.014	0.23	0.213	-0.196							
9. Preemptive help-seeking	0.719	0.628	0.669	0.598	-0.328	-0.354		-0.038	-0.214	-0.107	0.433	0.274		-0.034	0.074	0.277	-0.547	0.23
10. Familiarity with future opportunities	0.809	0.683	0.744	0.595	0.011	0.033		-0.017	0.072	0.17	-0.021	-0.032	0.056		0.296	0.234	0.276	0.225
11. Grit/persistence	0.778	0.715	0.772	0.66	0.276	0.18		-0.066	0.322	0.227	-0.15	0.332	-0.115	0.123		0.607	0.396	0.398
12. Leadership	0.82	0.683	0.77	0.616	0.105	0.067		-0.082	0.169	0.067	0.154	0.389	0.149	0.082	0.459		0.143	0.397
13. Self-efficacy/confidence	0.934	0.75	0.907	0.693	0.679	0.701		0.056	0.471	0.619	-0.621	0.368	-0.463	0.106	0.480	0.150		0.246
14. Social support	0.831	0.544	0.833	0.547	0.065	0.005		-0.019	0.214	0.146	-0.045	0.562	0.178	0.171	0.470	0.416	0.200	

toward CS (13), Behavioral intention (6), Beliefs about computer scientists (11), Computing identity (4), Current Experience (2), Asking for/receiving help (7), Familiarity with future opportunities (5), Grit/persistence (5), Leadership/teamwork (6), Self-efficacy/confidence (8), and Social support (8).

B. Results

1) Measurement Validity: To verify our factor structure, we conducted a CFA. During the analysis, items with low communality (< 0.4) [30] were removed. Due to a low number of fitting items, the *Beliefs about computer scientists* was subjected to an EFA and split into 3 factors: *Belief in the struggle of computer scientists*, *False belief in the effortless process of computer scientists* and *Belief in the insight of computer scientists*. A similar analysis split the *Asking for/receiving help* factor into 2: *Proactive help-seeking* and *Preemptive help-seeking*.

After this, CFA analysis was performed on the 14 factors. The model showed a reasonable fit: χ^2 (1517) = 3074.115, $p < .001$, with CFI = 0.938, TLI = 0.933, RMSEA = 0.055 (90% CI: [.052, .058]).

The factor fit metrics (Table II) showed that all AVEs except for *Belief in the struggle of computer scientists* were higher than 0.5. There was a lack of discriminant validity between the *Attitude toward CS* and *Behavioral intention* factors ($\sqrt{AVE} <$ the correlation with other factors, which can be found in the upper right of the diagonal in Table II). The factor loadings are presented in Table III. We could not provide a robust scale for the *Proactive help-seeking* factor.

2) Differences Between Groups: Analysis of demographic differences showed that the women in this sample had lower levels of **attitude towards CS** (1.032 v. 1.268), t(107.9) = 2.13, p = .035, d = 0.411), **behavioral intention** (0.293 v. 0.688), t(118.6) = 2.63, p = .010, d = 0.483), **current experience** (-0.833 v. -0.414), t(134.6) = 2.79, p = .006, d = 0.482), **false belief in the effortless process of computer scientists** (-1.127 v. -0.760), t(146.0) = 2.96, p = .004, d = 0.490), **computing identity** (-0.520 v. 0.013), t(118.8) = 3.60, $p < .001$, d = 0.660), and **self-efficacy/confidence** (0.335 v. 0.893), t(105.7) = 3.86, $p < .001$, d = 0.751) than the men. Conversely, they had a higher level of **belief in the struggle of computer scientists** (1.307 v. 0.942 on a scale from -2 to +2), t(167.9) = -4.25, $p < .001$, d = 0.656).

We also observed some significant differences in the factors between students in the CS1 and CS2 courses. CS2 students had higher levels of **attitudes towards CS** (1.479 v. 1.030), t(333.7) = -5.83, $p < .001$, d = 0.638), **intention** (1.059 v. 0.277), t(335.9) = -7.03, $p < .001$, d = 0.767), **computing identity** (0.355 v. -0.434), t(297.0) = -6.66, $p < .001$, d = 0.773), and **self-efficacy** (1.055 v. 0.564), t(330.4) = -4.66, $p < .001$, d = 0.513). Conversely, they reported a lower level of **perceived social support** (1.024 v. 1.201), t(297.3) = 2.14, p = 0.033, d = 0.248).

V. Study 3

A. Method

In the spring of 2023, we invited undergraduate students in CS courses from the same PWI to participate in our third survey. We expanded the students' grade level to include those enrolled in CS3 courses. A total of 459 participants took part in the survey. Out of these, the data of 9 students were removed for failing the attention-check question. Of the 450 used in the analysis, 354 (78.67%) identified as men, and 90 (20%) identified as women. 28 (6.22%) were Black. The number of participants enrolled in CS1, CS2, and CS3 courses were 149 (33.11%), 175 (38.89%), and 126 (28%), respectively.

We employed the same scale used in Study 2 to look for improvement in the scale's reliability and validity.

B. Results

1) Measurement Validity: Once the data was collected and cleaned, a CFA of the responses was carried out. Similar to our Study 2, while we were able to verify the convergent validity of the constructs in almost all the cases, that of *Proactive help-seeking* could not be verified due to an AVE less than 0.5. The Cronbach's alpha, AVE, and factor correlations are presented (in light blue) in Table II. The model demonstrated a reasonable fit: χ^2 (1517) = 3873.967, $p < .001$, CFI = 0.916, TLI = 0.908, RMSEA = 0.059 (90% CI: [.057, .061]).

Unlike Study 2, the current study found a robust factor for *Proactive help-seeking*, but the *Belief in the struggle of computer scientist* factor did not perform well. Loadings are presented in Table III.

2) Differences Between Groups: Similar to Study 2, the women in our sample reported lower **false belief in the effortless process of computer scientists** (-1.194 vs. -0.818 on a scale from -2 to +2), t(155.9) = 3.48, p = .001, d = 0.557), **computing identity** (-0.194 v. 0.115), t(131.0) = 2.27, p = .025, d = 0.397), **current experience** (-0.944 v.

TABLE III: Factor loadings. The new items added are highlighted with an asterisk. The greyed cells denote poorly fitting items. The green cells connote factors that have fit well over the studies and, therefore, can be used by other researchers

Factor	Item	Loadings	
		Study 2	Study 3
Attitude towards CS	I would voluntarily participate in additional computer science activities outside of classes, if provided the opportunity	0.767	0.744
	I expect that learning to use computer science skills will help me achieve my career goals	0.853	0.878
	I'll need a firm understanding of programming for my future work	0.873	0.924
	Knowing programming will help me earn a living	0.857	0.921
	It would make me happy to be recognized as an excellent student in computer science	0.790	0.784
	Being regarded as smart in computer science would be a great thing	0.729	0.722
	I study programming because I know how useful it is	0.780	0.802
	Programming is of no relevance to my life	-0.696	-0.777
	I see computer science as a subject I will rarely use in my daily life	-0.830	-0.781
	My goals do not require that I learn computer science skills	-0.898	-0.855
	I do not enjoy spending a lot of time writing programs	-0.730	-0.758
	Computer science is a worthwhile and necessary subject		
	I hope that in the future, I can find a career that does not require the use of computer science	-0.892	-0.785
Behavioral intention	Change to a different major or minor		
	Complete a CS major or minor	0.826	0.870
	Pursue a career after graduation where you directly apply the expertise and skills you will have acquired in your computing major or minor	0.989	0.978
	Acquire professional certifications related to your computing major or minor*	0.893	0.843
	Pursue a graduate degree in computing*	0.725	
	Consider a career where the knowledge and skills you will have acquired in your computing major or minor are not required*		
Belief in the struggle of computer scientists	Often struggle with fixing errors and bugs in their code		
	Sometimes have to spend a long time looking for simple errors, like typos	0.645	
	Rarely need to use Google or other resources to figure out how to solve their problems		
	Remember the syntax they need and rarely have to look it up		
	Often get help from others	0.671	
	Sometimes take longer than they expect to write their programs	0.732	
False belief in the effortless process of comp. scientists	Jump into writing code without having to think and plan much	0.778	0.840
	Program without having to think	0.884	0.774
Belief in the insight of computer scientists	Understand error messages relatively easily		
	Understand the programming task that they are given	0.779	0.857
	Know how to approach writing their programs	0.853	0.798
Computing identity	In general, being a computer scientist is an important part of my self-image	0.808	0.838
	I have a strong sense of belonging to the community of computer scientists	0.840	0.831
	I have come to think of myself as a "computer scientist"	0.852	0.810
	Overall, I feel like I belong in computer science	0.994	0.973
Current Experience	The course work in my current course required more programming experience than I had	0.923	0.875
	The course work in my current course required less programming experience than I had	-0.788	-0.771
Proactive help-seeking	I am not afraid to ask for help		0.638
	I am comfortable asking my professors or other instructors (TAs, other CS lecturers, etc.) for help with my course work		0.633
	This course provides me with communication/feedback (e-mails, online course posts, written communication) that is easy to access		0.796
Preemptive help-seeking	I ask for help more often than my classmates	0.821	0.853
	I ask for help as soon as something is unclear to me	0.763	0.684
Familiarity with future opportunities	Applying to graduate school	0.824	0.833
	The graduate school experience	0.855	0.865
	Computer science industry positions	0.912	0.636
	Graduate school funding opportunities	0.780	0.814
	Technology entrepreneurship opportunities	0.753	0.683
Grit / Persistence	I am a hard worker	0.819	0.849
	I am diligent	0.801	0.856
	I can overcome obstacles to complete my tasks	0.912	0.799
	I am a procrastinator		
	I give up when problems seem too difficult		-0.739
Leadership	I know how to cooperate effectively as a member of a team	0.846	0.843
	I have high confidence in my ability to function as part of a team	0.926	0.852
	I know a lot about what it takes to be a good leader	0.724	
	I value the contributions of my team members*		0.775
	I treat my team members as equal members of the team*		0.719
	I am good at communicating with my team members*	0.796	0.724
Self-efficacy/confidence	I am confident I will do well on computer science labs and projects	0.869	0.794
	Even when I work hard at it, programming tends to be unusually hard for me	-0.775	-0.713
	I am sure I can understand computer science	0.946	0.825
	I am confident in my problem solving ability	0.760	
	I am confident in my ability to meet unexpected programming challenges with success	0.878	0.830
	I am confident in my ability to complete a major in computer science	0.911	0.960
	I think I could handle difficult computing problems	0.881	0.858
	I'm not good at computing	-0.893	-0.828
Social support	I have friends at school that support and care about me	0.659	0.674
	I receive enough support and resources to meet the unique challenges of the school year from my school administration	0.779	0.832
	Students at my school have the support and resources they need to be successful with their learning	0.720	0.745
	In my major, there is at least one professor or lecturer who listens to what I have to say	0.633	0.639
	When I need suggestions on how to deal with a school problem, I know someone I can turn to*	0.765	0.799
	I feel there is no one I can share my school problems with*	-0.716	-0.730
	There are people that I trust to help solve my school problems*	0.865	0.741
	There is someone who takes pride in my accomplishments*		

-0.561), $t(167.2) = 3.274$, $p = .001$, $d = 0.506$), and **self-efficacy/confidence** (0.340 v. 0.870), $t(129.1) = 4.77$, $p < .001$, $d = 0.840$) levels than the men. On the other hand, they reported higher levels of **preemptive help-seeking** (-0.276 v. -0.583), $t(130.6) = -2.34$, $p = .021$, $d = 0.409$), **grit/persistence** (1.428 v. 1.266), $t(137.8) = -2.19$, $p = .030$, $d = 0.373$), and **leadership** (1.680 v. 1.572), $t(159.3) = -2.15$, $p = .033$, $d = 0.342$).

In terms of race, we found that our small sample of Black students had a higher level of **preemptive help-seeking** compared to the non-Black students: (-0.036 v. -0.536), $t(30.4) = 2.20$, $p = .036$, $d = 0.797$).

There were also some significant differences in the factors among students in CS1, CS2, and CS3. For instance, among the three groups, there was a difference in their intention to continue their CS trajectory: $F(2, 447) = 27.3$, $p < .001$. Post-hoc tests revealed significant differences between the CS1 and CS2 as well as the CS1 and the CS3 groups. Specifically, both CS2 (1.354 v. 0.651, $p < .001$) and CS3 (1.495 v. 0.651, $p < .001$) had higher levels of **intention** than the CS1 students. The same pattern was observed in their **attitude towards CS**, $F(2, 447) = 16.8$, $p < .001$, and reported levels of **computing identity**, $F(2, 447) = 16.8$, $p < .001$ and **self-efficacy/confidence**, $F(2, 447) = 12.9$ $p < .001$.

VI. Discussion

Our study revealed significant demographic disparities in the realm of computing sciences. Interestingly, in addition to differences between genders and among students in the CS courses, Black students demonstrated a higher **familiarity with future opportunities** in the field than non-Black students, suggesting strong career awareness and ambition. This trend was especially pronounced among HBCU students compared to Black students at the PWI. However, despite these optimistic indicators, Black students also exhibited higher levels of **preemptive help-seeking** behavior, which could reflect an acute awareness of challenges or perceived barriers in CS programs.

The findings of this study emphasize the importance of acknowledging and addressing the unique experiences and challenges faced by diverse groups within the field of computer science. For instance, given the interest of Black students in future opportunities, integrating career development into CS education can help them navigate and prepare for the professional world effectively. On the other hand, their preemptive help-seeking behavior indicates a requirement for robust support systems. This may include mentorship, tutoring, and resources tailored to address the unique challenges Black students face in CS.

One limitation identified in this study is the small sample size of Black students. In future research, we plan to expand our data-gathering efforts by including a larger sample of Black students from the two HBCUs alongside our ongoing data collection from the PWI.

VII. Conclusion

This research is part of a longitudinal project that seeks to identify inclusive strategies for success in CS programs and investigate ways to increase Black students' retention, graduation, and post-graduation success. It features a collaboration between two Historically Black Colleges and Universities (HBCUs) and a Primarily White Institution (PWI), at various levels on the Carnegie Classification of Institutions of Higher Education. We plan to conduct more studies with data from the three institutions to evaluate and refine the scales presented in this paper. In the meantime, we summarize the lessons learned so far in our scale development process.

Researchers often create scales on-the-fly to address the specific needs of their study. This *ad hoc* approach can jeopardize measurement validity and the generalizability of the study results. In the current work, we set out to develop a more robust set of scales that can be used to carefully measure the subjective factors impacting the retention of URM students in the undergraduate CS curriculum, and find it to be a fruitful but iterative endeavor. Indeed, only through this multi-stage, persistent effort will one be able to create a robust measurement instrument that can truly serve as a valuable resource for researchers, educators, and policymakers.

During our three studies, we revisited initial assumptions about what factors motivate and limit students' success, revised questions aimed at measuring these factors, and conducted multiple iterations of testing, experiencing different levels of confidence in how well factors worked. Factors such as *Belief in the struggle of computer scientists* and *proactive help-seeking*, are not so well fitted to the current model.

In the case of the *Attitude toward CS* and *Behavioral intention* constructs, though they lacked discriminant validity, we note that intention is a behavioral construct, while attitude, as the name implies, is an attitudinal construct. Mixing constructs with different psychological roots tends to result in a lack of robustness, hence we hesitate to combine them.

For now, researchers will have to choose between our study 1 scales, which showed stronger discriminant validity and our study 2+3 scales, which have a more robust theoretical base because of their increased number of items. In the future, we plan to further improve upon these scales by adding/adjusting items and by iterating on their composition so as to guarantee both. However, the following factors fit well throughout our studies, so they are robust enough to be used in future studies:

- *Computing identity*
- *Current experience*
- *Familiarity with future opportunities*
- *Self-efficacy/confidence*
- *Social support*

Beyond establishing the reliability and validity of our scales, we have demonstrated that they can be used to uncover differences between different groups of students. This is of practical importance for the use of our scales in CS education research, especially in research on equity and inclusion. We present our current scale development efforts as a useful demonstration and guideline. We intend for this research to provide insights into a thorough and repeatable scale development process that can be used by the computing education research community, as well as the science education community at large.

REFERENCES

[1] L. J. Charleston, P. L. George, J. F. L. Jackson, J. Berhanu, and M. H. Amechi, 'Navigating underrepresented STEM spaces: Experiences of Black women in US computing science higher education programs who actualize success', Journal of Diversity in Higher Education, vol. 7, no. 3, p. 166, 2014.

[2] O. Osho et al., 'Societal Factors that Impact Retention and Graduation of Underrepresented Computer Science Undergraduates', in Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 2, 2022, pp. 1399–1399.

[3] M. Kordaki and I. Berdousis, 'Identifying Barriers For Women Participation In Computer Science', Pro Edu. International Journal of Educational Sciences, vol. 2, no. 2, pp. 5–20, 2020.

[4] M. N. Giannakos, I. O. Pappas, L. Jaccheri, and D. G. Sampson, 'Understanding student retention in computer science education: The role of environment, gains, barriers and usefulness', Education and Information Technologies, vol. 22, pp. 2365–2382, 2017.

[5] C. Kennedy, 'An Exploration of Student Reasoning about Undergraduate Computer Science Concepts: An Active Learning Technique to Address Misconceptions', Clemson University, 2020.

[6] C. Schulte and M. Knobelsdorf, 'Attitudes towards computer science-computing experiences as a starting point and barrier to computer science', in Proceedings of the third international workshop on Computing education research, 2007, pp. 27–38.

[7] B. Brinkman and A. Diekman, 'Applying the communal goal congruity perspective to enhance diversity and inclusion in undergraduate computing degrees', in Proceedings of the 47th ACM technical symposium on computing science education, 2016, pp. 102–107.

[8] M. M. Hossain and M. G Robinson, 'How to motivate US students to pursue STEM (science, technology, engineering and mathematics) careers', Online Submission, 2012.

[9] S. Olson and D. G. Riordan, 'Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Report to the president', Executive Office of the President, 2012.

[10] C.-S. G. Davis and C. J. Finelli, 'Diversity and retention in engineering', 2007.

[11] M. R. H. Roberts, T. McGill, and T. Koppi, 'What students are telling us about why they left their ICT course', Innovation in Teaching and Learning in Information and Computer Sciences, vol. 10, no. 3, pp. 68–83, 2011.

[12] N. A. Buzzetto-More, O. Ukoha, and N. Rustagi, 'Unlocking the barriers to women and minorities in computer science and information systems studies: Results from a multi-methodolical study conducted at two minority serving institutions', Journal of Information Technology Education: Research, vol. 9, no. 1, pp. 115–131, 2010.

[13] C. Kennedy and E. T. Kraemer, 'What are they thinking? Eliciting student reasoning about troublesome concepts in introductory computer science', in Proceedings of the 18th Koli Calling International Conference on Computing Education Research, 2018, pp. 1–10.

[14] K. Goldman et al., 'Identifying important and difficult concepts in introductory computing courses using a delphi process', in Proceedings of the 39th SIGCSE technical symposium on Computer science education, 2008, pp. 256–260.

[15] U. Bunz, 'The computer-email-web (CEW) fluency scale-development and validation', International journal of human-computer interaction, vol. 17, no. 4, pp. 479–506, 2004.

[16] A. Lishinski, S. Narvaiz, and J. M. Rosenberg, 'Self-efficacy, Interest, and Belongingness–URM Students' Momentary Experiences in CS1', in Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1, 2022, pp. 44–60.

[17] M. Biggers, A. Brauer, and T. Yilmaz, 'Student perceptions of computer science: a retention study comparing graduating seniors with cs leavers', Acm sigcse bulletin, vol. 40, no. 1, pp. 402–406, 2008.

[18] R. A. Mickelson, I. Mikkelsen, M. Dorodchi, B. Cukic, and T. Horn, 'Fostering greater persistence among underserved computer science undergraduates: a descriptive study of the I-PASS project', Journal of College Student Retention: Research, Theory & Practice, p. 15210251221086928, 2022.

[19] N. A. Buzzetto-More, O. Ukoha, and N. Rustagi, 'Unlocking the barriers to women and minorities in computer science and information systems studies: Results from a multi-methodolical study conducted at two minority-serving institutions', Journal of Information Technology Education: Research, vol. 9, no. 1, pp. 115–131, 2010.

[20] C. Mooney and B. A. Becker, 'Sense of belonging: The intersectionality of self-identified minority status and gender in undergraduate computer science students', in United Kingdom & Ireland Computing Education Research conference., 2020, pp. 24–30.

[21] N. Schauffel, I. Schmidt, H. Peiffer, and T. Ellwart, 'Self-concept related to information and communication technology: Scale development and validation', Computers in Human Behavior Reports, vol. 4, p. 100149, 2021.

[22] D. Harrington, Confirmatory factor analysis. Oxford university press, 2009.

[23] K. M. Lewis and P. Hepburn, 'Open card sorting and factor analysis: a usability case study', The electronic library, vol. 28, no. 3, pp. 401–416, 2010.

[24] L. R. Fabrigar and D. T. Wegener, Exploratory factor analysis. Oxford University Press, 2011.

[25] Y. U. N. Zhong-Lin, H. Ban-Ban, and T. Dan-Dan, 'Preliminary work for modeling questionnaire data', Journal of Psychological Science, vol. 41, no. 1, p. 204, 2018.

[26] L. M. Lower, T. J. Newman, and D. Anderson-Butcher, 'Validity and reliability of the teamwork scale for youth', Research on Social Work Practice, vol. 27, no. 6, pp. 716–725, 2017.

[27] S. Cohen, R. Mermelstein, T. Kamarck, and H. M. Hoberman, 'Measuring the functional components of social support', in Social support: Theory, research and applications, Springer, 1985, pp. 73–94.

[28] S. Cohen and H. M. Hoberman, 'Positive events and social supports as buffers of life change stress 1', Journal of applied social psychology, vol. 13, no. 2, pp. 99–125, 1983.

[29] C. E. Osgood, 'Semantic differential technique in the comparative study of cultures', American Anthropologist, vol. 66, no. 3, pp. 171–200, 1964.

[30] M. Iqbal, 'A study of the impacts of social background on the learning experiences and approaches of students in architectural design education in Pakistan', Cardiff University, 2020.