Confidence, Interest, and Gender Perception in non-Computer Science Majors: an Instrument Re-validation Study

Sara Hooshangi
Computer Science Department
Virginia Tech
Blacksburg, VA, USA
shoosh@vt.edu

Khushi Parajuli
Computer Science Department
Virginia Tech
Blacksburg, VA, USA
kparajuli@vt.edu

Brandi A. Weiss
Educational Leadership Department
George Washington University
Washington D.C., USA
weissba@gwu.edu

Abstract

To broaden participation in the computer science (CS) field and its workforce, it is important to consider how students from non-CS majors enter the field at various points along the educational pipeline. Gaining insight into these students' attitudes and interests toward CS requires a validated, reliable instrument that can capture the factors influencing their perceptions. While several tools have been developed to measure motivation, attitudes, knowledge, and self-efficacy in CS, few are specifically designed to focus on non-CS majors who may hold peripheral or emerging interests in the discipline. In this study, exploratory factor analysis was used to re-validate the Engineering Students' Attitudes towards CS survey initially created by Hoegh and Moskal using a population of non-CS majors. Results indicated that a 1-factor solution best fits the data for the Interest, Confidence, and Gender Equality Perceptions (GEP) constructs. Unique to this study, is support for a shortened 5-item GEP subscale. Results showed that the 5-item GEP performed as well as (and at times better than) the 10-item GEP. Based on these results, we recommend researchers wishing to examine Gender Equality Perceptions use a shortened version of the subscale utilizing only the 5 positively worded items. As a secondary interest of the work, results indicated women were nearly a full standard deviation higher on GEP subscales (Cohen's d = .961 and .837). This is considered a large effect size in social science research and indicates women had higher ratings of gender equality in CS than men did.

CCS Concepts

• **Social and professional topics → Computing education**.

Keywords

Confidence, interest, gender perception, instrument re-validation

ACM Reference Format:
Sara Hooshangi, Khushi Parajuli, and Brandi A. Weiss. 2025. Confidence, Interest, and Gender Perception in non-Computer Science Majors: an Instrument Re-validation Study. In *Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2025), June 27-July 2, 2025, Nijmegen, Netherlands.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3724363.3729042

ITiCSE 2025, Nijmegen, Netherlands

ACM ISBN 979-8-4007-1567-9/2025/06
https://doi.org/10.1145/3724363.3729042

1 Introduction

Recent studies have identified several barriers to degree completion and career entry in CS including stereotypes, biases, classroom cultures [12, 40], lack computer self-efficacy, and a wavering sense of belonging [20, 34]. On the other hand, confidence [29], self-efficacy [7] and pre-college experiences [47, 50] have been found to positively affect the student experience in computer science (CS). Students who have positive attitudes towards computing are more likely to persist in computing [15]. Additionally, researchers have attributed internal factors such as attitudes and perceptions towards computer science at various education levels as important factors for shaping student's trajectories [15, 48]. Understanding students' perceptions and attitudes towards the CS field may help mitigate the impact of negative stereotypes and encourage a diverse range of individuals to explore Computer Science.

Valid and reliable instruments measuring attitudes in CS are essential to understanding how these constructs impact outcomes. Several instruments have been developed to measure motivation, attitudes, knowledge, and student self-efficacy in CS [13, 15, 23–25, 33, 48]. While many of these surveys focus on attitudes at the undergraduate or secondary school level, few are focused on students that are not part of the CS pipeline (i.e. non-CS majors) [27].

The purpose of the current study was two-fold. The primary goal of this study was to re-validate an instrument [26] measuring the confidence, interest, and gender perceptions of students towards CS for a population of non-CS majors enrolled in introductory CS courses. The secondary goal of the current study was to use the re-validated instrument to investigate differences in confidence, interest, and gender perceptions of CS coursework across gender, ethnicity, and college majors of non-CS students using inferential statistics. Previously, this instrument was validated using a sample of freshman engineering students enrolled in a Calculus II course, a required mathematics course students for science and engineering students that was typically completed before selecting a major [26, 35], and for students in high school who participated in summer workshops [22]. The validity and reliability of measures are always specific to the population of interest. An instrument may be valid and reliable for one population but not for another. Because the population of interest in the current study differed from that of the original study, re-validation was necessary. In re-validating this instrument, we provide validity and reliability evidence for using the measure in a different population, providing evidence that the instrument measured what it purported to measure and that the items were internally consistent [31].

2 Prior Work

The recent literature on STEM education has shed light on various aspects of retention challenges in higher education. Blackburn's [4] comprehensive review highlights the status of women in STEM from 2007 to 2017, emphasizing the need for ongoing research and interventions. Giannakos et al. [19] examined the factors that influenced student retention in computer science education, identifying the crucial role of the learning environment, perceived gains, barriers, and perceived usefulness of the field. Tellhed et al. [44] contributed insights into the factors shaping adolescents' first educational choices, emphasizing the role of ability beliefs and acetic versus communal career goals, providing explanations for gender imbalances. Other studies have focused on the persistence and retention of students already in the computing pipeline [4, 8, 42]. In this work, we were interested in understanding the attitude and perception of non-CS majors in introductory CS courses about the computing field. As such, we began our investigation by looking at the literature around attitude in CS and examined existing instruments that have been developed and validated to measure relevant constructs. We will start by doing a looking at existing surveys and then focus on the chosen instrument for this study.

2.1 Engineering Attitude Surveys

Various surveys have been designed to assess constructs related to engineering education across different education levels. Engineering Attitude Survey [36] examined constructs such as transfer, interest, problem-solving, confidence, attitude, and motivation. The Engineering Motivation Survey [6] focuses on undergraduate students and measures constructs such as attainment, interest, cost, utility value, and expectation of success. The Freshman Engineering Attitude Survey [21] targets undergraduate students, exploring communication skills, knowledge integration, lifelong learning, team expectations, and technical skills. High school students are the focus of the High School Students' Attitude to Engineering Scale [25], which evaluates confidence, career aspirations, self-efficacy, academic history, knowledge, and demographics. The Information Technology Attitude Survey [18] caters to both high school and undergraduate levels, measuring confidence, interest, gender, usefulness, and professionalism. Middle school students are considered in the Middle School Students' Attitude to Mathematics, Science, and Engineering survey [24], covering attitude, knowledge, academic performance, and engineering discussions. The Pittsburgh Engineering Attitude Scale-Revised [23] targets undergraduates, exploring general impressions, financial influences, societal contributions, social prestige, enjoyment, career aspects, and parental pressure. The Student Attitude Survey [33], administered at the undergraduate level, investigates problem-solving, technical roles, financial issues, ethics, environmental impact, sustainability, and diversity. The diversity of these surveys provides a comprehensive view of students' attitudes and perceptions in various educational contexts in the context of STEM education, however they may not address or measure constructs that are more prevalent in CS.

2.2 CS Attitudes Surveys

While engineering attitude surveys can provide a more general perspective of students attitude towards engineering field as a whole, it is not clear if all the measured factors will translate into the computing field, especially since many CS departments do not reside within a school of engineering. Researchers have developed several specific CS surveys to focus on the students in the CS pipeline. Ramalingam and Wiedenbeck investigated the relationship between student attitudes and self-efficacy and learning how to program [39] by developing the Computer Programming Self-Efficacy Scale in 1998. Wiebe et al. [49] created the Computing Attitude Survey based on a previously developed mathematics attitudes scale [17], to examine confidence, attitude, male dominance, usefulness and motivation. Hoegh and Moskal developed the Engineering Students' Attitudes toward CS [26] survey which assessed confidence, interest, gender, usefulness, and professionalism. A group of researchers adapted the Colorado Learning Attitudes about Science Survey (CLASS) [41] to design a new instrument called Computing Attitudes Survey (CAS) [15], to measure novice to expert-like perceptions about Computer Science and problem solving skills. The survey was later validated with students in an introductory CS and it is currently available at its fourth version [13].

Several other validated instruments have focused on other aspect of the computing field such computational thinking skills scale [32], programming self-efficacy scale [45] and the development of a Computer Science Attitudes Scale for middle school students (MG-CS attitudes) [38]. A recent study has compared student responses in a CS0 and CS1 course and has shown increased in confidence in CS0 students and an overall initial lack of confidence and fear among non-majors enrolled in introductory CS classes [27]. However this study does not use a validated survey and the results are based on open ended questions and a single Likert-type question that is asked over the course of a semester to measure confidence over time. A recent 2020 comprehensive literature review of the existing instrument in CS education has revealed that many of the existing surveys are not fully validated [51] and more work is needed to provide transparent use of surveys and reproducible practices.

For our study, we decided to use the Hoegh and Moskal [26] survey because of the similarity of the population under the study which would allow us to re-validate the survey and use it to investigate the attitude of non-CS majors. While their study specifically focused on freshmen science and engineering students who have not declared a major yet, the sample population for the current study is broader and encompasses any student outside the realm of CS who completes an introductory CS course (not counted towards the CS major) and in most likelihood will not become a CS major. In the following section, we will give details of the survey.

2.3 Engineering Students' Attitudes Toward Computer Science Survey- ATCS

Hoegh and Moskal's Engineering Students' Attitudes toward CS [26] survey, called ATCS for the remainder of this paper, assesses confidence, interest, gender, usefulness, and professionalism. In a 2009 paper, Hoegh and Moskal [26] investigated the attitudes of freshmen science and engineering students toward CS to shed light on factors that may influence student choice of major and students' perceptions of computer science as a male-dominated field. The authors developed an attitude instrument with five different constructs: Confidence(C), Interest(I), Gender(G), Usefulness(U), and Professional(P). The study reveals a diversity of attitudes among science and engineering freshmen students in a Calculus II class regarding CS. It identifies a spectrum of sentiments, ranging from enthusiastic interest to potential aversion, suggesting that students' perceptions of CS are not uniform within these disciplines. The authors find that students' prior experiences with CS significantly shape their attitudes. Those who have prior exposure through coursework or other activities generally express more positive attitudes, highlighting the impact of early exposure on shaping perceptions. The paper uncovers gender differences in attitudes toward CS. The paper further discussed how stereotypes and societal expectations may contribute to these differences, and emphasizes the importance of addressing gender-related barriers to foster more inclusive attitudes among all students. The authors identified challenges in altering negative perceptions. Addressing these challenges requires not only curricular adjustments but also a concerted effort to create inclusive learning environments that dispel stereotypes and encourage diverse student participation. Hoegh and Moskal emphasized the necessity of adopting inclusive practices within CS education. This includes promoting diversity, addressing gender disparities, and creating an environment that welcomes students from various backgrounds. In summary, Hoegh and Moskal's study provided valuable insights into the attitudes of freshmen science and engineering students toward CS. The findings help understand the factors influencing students' perceptions and suggest strategies for fostering a more inclusive and positive environment for learning CS within these disciplines so that a more diverse student population enters the CS field.

3 Conceptual Framework and Research Questions

This study is guided by Stephens, Markus and Fryberg's Sociocultural Self Model of Behavior (SSMB) [43]. Building on theories spanning health psychology, social and cultural psychology, and sociology [2, 5] the SSMB combines individual and structural models and is based on the argument these two constructs are inseparable and therefore must be analyzed together. The SSMB can be used to investigate micro- and macro-level factors that contribute to attitude towards the CS field. Student choices are made within a larger social and structural environment and barriers faced by non-CS students may be higher than those of their peers in general engineering or CS field, and their entry into the major may be more challenging. Social and structural constructs work together to understand factors that persuade or dissuade such individuals. Guided by this theory, we aim to examine the attitude of non-CS majors. Our research questions are as follows:

- RQ1: What is the underlying factor structure of the ATCS for undergraduate non-CS majors, and is it reliable?
- RQ2: Do men and women differ in their confidence, interest, and gender equality perceptions in CS?

4 Methodology

4.1 Population and Participants

The population of interest for the current study was undergraduate non-computer science (CS) majors enrolled in an introductory CS course at Virginia Tech which is a large R1 institution in the United States. The CS department at our institution offers 5 introductory CS courses designed for non-majors: 1) Introduction to Computational Thinking, 2) Introduction to Programming in C, 3) Introduction to Programming in Java, 4) Introduction to Programming in Python, and 5) Intermediate Programming in Python. Each semester, approximately 400 undergraduate students enroll in 1 or more of these 5 courses to fulfill a general education requirement or for personal interest. These courses were intentionally selected to recruit students for the current study because they contain students from the target population of non-CS students.

Data were collected from 158 students enrolled in at least one of these courses during the Spring 2023 semester. After data were cleaned, 24 participants were removed due to missing identification numbers, duplicate IP addresses, or total survey response times of <1 minute. This resulted in a sample of 134 participants. Of these, 6-13% did not respond to some items at the end of the survey, depending on the item. Listwise deletion was utilized for each analysis to retain as many participants as possible, and because data were missing not at random. This resulted in a slightly smaller sample for some analyses, noted where appropriate in the results. The demographics for the 134 participants are shown in Table 1.

Table 1: Survey Respondents Demographics

Item	Categories	# of Respondents (%)
Gender	Male	56 (42%)
	Female	49 (36%)
	Non-binary	4 (3%)
	Not-Reported	25 (19%)
Race /Ethnicity	White	70 (52%)
	URM	40 (30%)
	Not-Reported	24 (18%)
Standing	Freshman	39 (29%)
	Sophomore	40 (30%)
	Junior	11 (8%)
	Senior	19 (14%)
	Not-Reported	25 (19%)
Total		134 (100%)

4.2 Measures

The current study utilized Hoegh and Moskal's (2009) "Attitudes Toward CS" (ATCS) instrument [26]. This instrument was selected because, 1) it measured attitude constructs relevant to the current research study, and 2) it was developed using a sample of first-year non-CS majors, which was similar (albeit, not-identical) to the target population of the current study. The Hoegh and Moskal (2009) instrument was designed to measure five constructs: "Confidence", "Interest", "Gender", "Usefulness", and "Professional" in CS education. In the current study, we selectively focused on the Confidence, Interest, and Gender subscales of this instrument to better reflect the constructs of interest. Confidence is defined as students' confidence in acquiring skills related to CS and was measured by 8 items; Interest is defined as students' interests in CS and was measured by 10 items; and what Hoegh and Moskal (2009) refer to as gender for clarity we renamed as Gender Equality Perceptions (GEP) which is defined as students' views on CS being a male-dominated field and was measured by 10 items.

All 28 items were measured on a 1 to 5 Likert-type response scale, where 1 represented "strongly disagree" and 5 represented "strongly agree". Four items, five items, and five items were negatively worded on the Confidence, Interest, and GEP subscales, respectively. Negatively worded items were not reverse-coded before examining the measurement structure via factor analyses but were later reverse-coded for reliability analyses and to compute composite scores for inferential analyses.

4.3 Procedure

Data were collected using a survey created in Question Pro. Participants were recruited using e-brochures sent by the instructors of the five introductory CS courses. In addition to the 28 items from Hoegh and Moskal's (2009) instrument, demographic information was collected at the end of the survey. Students who completed the survey were entered into a drawing for Amazon gift cards. The survey was administered towards the end of the semester, and aimed to capture a comprehensive range of experiences and attitude shifts.

4.4 Statistical Analyses

Data were screened and analyzed using Python 3.12 and the Statistical Package for Social Sciences (SPSS) version 29. The original Hoegh and Moskal's (2009) instrument was developed using first-year students enrolled in a school primarily focused on science and engineering. Because the current study aimed to collect data from students across a more diverse group of majors, it was necessary to re-validate the instrument before examining gender differences. To examine the research questions, the current study utilized exploratory factor analysis (EFA) and independent t-tests.

To determine the factorability of data, 3 methods were utilized: Bivariate Pearson correlation coefficients, the Kaiser Meyer Olkin (KMO) test [30], and Bartlett's test of Sphericity [3]. All three methods indicated data were appropriate for factor analysis (i.e., correlations for items within the same subscale were $>|.30|$, KMO values exceeded 0.7, and Bartlett's test was statistically significant).

To examine the dimensionality of the three subscales (Confidence, Interest, and GEP), exploratory factor analysis (EFA) was conducted using Principal axis factoring (PAF) separately for each subscale. PAF, which minimizes the residuals between the observed and model-implied correlations, was used to reflect the latent nature of the subscales. Bandalos and Finney [1]state researchers should always use multiple extraction techniques to determine the number of factors to extract. In the current study, the following extraction methods were utilized: eigenvalues greater than 1, scree plot, Cattell-Nelson-Gorsuch objective scree plot, Velicer's minimum average partial (MAP) procedure [46], and Horn's parallel analysis [28]. O'connor's programs were used to conduct Velicer's MAP and Horn's parallel analyses [37]. We also considered the size of communalities extracted by the solution, the size and pattern of factor loadings, reliability estimates per factor, as well as theory and previous research. Multiple factor solutions were compared, and when a multi-factor solution was extracted Oblimin rotation was used to allow factors to be oblique (correlated). Oblique rotation was justified in this context as the underlying dimensions of measures in social or behavioral sciences are typically correlated [1].

For the final factor solution, composite scores were created for each factor by reverse coding negatively worded items, then averaging subscale items within a factor. This placed composite scores on a 1 to 5 scale, where a 5 reflected high attitudes on the factor.

To examine gender differences on each factor, independent t-tests were used. The following assumptions were examined for t-tests: normal distribution (Shapiro-Wilk) and Levene's test for homogeneity of variances (HOV). For analyses in which the HOV assumption was not met, a Welch-Aspen t-test was examined. A Bonferonni-adjusted alpha value of .0125 was used to determine the statistical significance of t-tests, keeping familywise type I error at .05. Cohen's d values [10] are reported as standardized effect sizes using benchmarks of 0.2 as small, 0.5 as medium, and 0.8 as large.

5 Results

5.1 RQ1: What is the underlying factor structure of the ATCS for undergraduate non-CS majors, and is it reliable?

To examine the dimensionality of the three subscales (Confidence, Interest, and GEP), EFA sing PAF was conducted separately for each subscale. Bandalos and Finney [1] state researchers should always use multiple extraction techniques to determine the number of factors to extract. Thus, five extraction methods were utilized in the current study: the following extraction methods were utilized: scree plot, Cattell-Nelson-Gorsuch objective scree plot, eigenvalues greater than 1, Velicer's minimum average partial (MAP) procedure [46], and Horn's parallel analysis [28].

The scree plot techniques are based on visual inspection of the eigenvalues to determine the number of factors to extract. Based on the scree plot and Cattell-Nelson-Goruch's (CNG) objective scree plot, the Confidence and Interest subscales were best represented by 1 factor, whereas GEP could be represented by 1, 2, or 3 factors. The eigenvalues greater than 1 technique is a mathematically-based method frequently used in EFA as it is the default in many statistical software programs. Velicer's MAP procedure and Horn's parallel analysis are statistically-based methods that perform well in simulation studies [1]. Based on eigenvalues greater than 1, Velicer's MAP, and Horn's parallel analysis, 1 factor best represented Confidence and Interest, whereas 2 factors best represented GEP.

To determine the number of factors to retain we also took into consideration the size of communalities extracted by the solution, the size and pattern of factor loadings, reliability estimates per factor, as well as theory and previous research. In summary, for the Confidence and Interest subscales, all extraction techniques suggested a one-factor solution, which was consistent with Hoegh and Moskal's (2009) findings and theory.

For the GEP subscale, the extraction methods were inconsistent with one another, thus 1-factor, 2-factor, and 3-factor solutions were examined using oblique Oblimin rotation for multi-factor solutions. The 3-factor solution didn't make theoretical sense and resulted in too few items per factor. In all 3 possible solutions, the first factor extracted contained the 5 positively worded items from the GEP subscale (i.e., items GEP3, GEP4, GEP8, GEP9, and GEP10). For the 2-factor solution, the 5 negatively worded GEP items loaded on

Table 2: Pattern coefficients for each subscale

Confidence		Interest		GEP-10		GEP-5
Item	Loading	Item	Loading	Item	Loading	Loading
C1	0.730	I1*	-0.682	GP1*	-0.741	
C2*	-0.651	I2*	-0.661	GP2*	-0.754	
C3*	-0.671	I3	0.825	GP3	0.683	0.726
C4	0.764	I4*	-0.774	GP4	0.794	0.902
C5	0.659	I5	0.742	GP5*	-0.650	
C6	0.778	I6*	-0.736	GP6*	-0.764	
C7*	-0.809	I7	0.680	GP7*	-0.609	
C8*	-0.771	I8*	-0.799	GP8	0.772	0.791
		I9	0.660	GP9	0.870	0.900
		I10	0.763	GP10	0.818	0.902
n	134		124		116	116
α	0.896		0.919		0.919	0.919

*negatively worded items; n =sample size;
GP=Gender Perceptions; α = Cronbach's coefficient alpha;
Blank cells indicate the item was not included in the analysis.

the second factor. This finding indicates a potential method effect in which negatively worded items have shared variance due to the wording rather than the data representing two unique constructs. Furthermore, the 2-factor solution exhibited absence of simple structure (with potential cross-loadings), and omitting 1 or more of the problematic items (GP5, GP7, and GP1) resulted in pattern coefficients over 1, an ultra-Heywood case, and Cronbach's coeffcient alpha increasing if these itesm were omitted.

We determined a 1-factor solution best fit the data for the GEP subscale and was consistent with theory and findings of Hoegh and Moskal (2009). Of note, we report two potential solutions for this GEP subscale, one in which all 10 items load onto a single factor (labeled GEP10) and another with only the 5 positively worded items included (labeled GEP5). Factor analysis results from this point on reflect the 1-factor solution all subscales (Confidence, Interest, and GEP). Results are presented for both the 10-item GEP and the 5-item GEP.

Table 2 contains the pattern coefficients for each subscale. Of note, in a 1-factor solution, rotation is unnecessary and both the pattern and structure coefficients are analogous. All pattern coefficients were greater than |.60| indicating all items within a subscale loaded onto a single factor. Table 2 also contains Cronbach's coefficient α values representing the internal consistency reliability for each factor solution. These were all high and approximately .90 indicating each factor was highly reliable. Alpha if any item was deleted from the Confidence or Interest subscales, resulted in decreased reliability estimates, indicating each item contributed to the overall reliability of the scale. For both the GEP-10 and GEP-5 subscales, coefficient alpha was .919. For the GEP5, alpha decreased if any items were deleted. Of note, if item GP7 was deleted from the 10-item GEP factor, reliability increased slightly to .921, indicating this item may be detracting from the overall reliability of the GEP-10 subscale.

Furthermore, nearly all communalities fell within the desirable 0.40 to 0.70 range, indicating each variable was explained well by the factor solution [11]. Of note, the communality for item GP7 was a little low (.371) on the 10-item GEP, indicating the variable may not have been explained well by the factor solution. Because this was a negatively worded item, it was excluded in the 5-item GEP, and thus not problematic. Researchers typically want to explain at least 50% variance from a factor solution. A one-factor solution explained 59.16%, 58.42%, 60.32%, and 77.12% of the variance in the Confidence, Interest, GEP-10, and GEP-5 subscales, respectively.

Table 3: Bivariate Pearson correlation coefficients

	Confidence	Interest	GEP-10	GEP-5
Confidence	1			
Interest	.704*	1		
GEP-10	.227*	0.14	1	
GEP-5	.250*	0.16	.917*	1

*p<.05 (two-tailed).
GEP-10 = Gender Equality Perceptions 10 items scale;
GEP-5 = Gender Equality Perceptions 5 item scale.

For the final factor solution, composite scores were created for each factor by reverse coding negatively worded items, then averaging subscale items within a factor. This placed composite scores on a 1 to 5 scale, where a 5 reflected high attitudes on the factor. Table 3. contains the bivariate Pearson correlations between the factors. Confidence and Interest were highly related (r=.704), whereas both were weakly related to GEP (r ranging from .14 to .25). Unsurprisingly, the 10- and 5-item GEP subscales were highly related (r=.917).

5.2 RQ2: Do men and women differ in their confidence, interest, and gender equality perceptions in CS?

Descriptive statistics for gender comparison are presented in Table 4. Independent t-tests were conducted to compare men and women on Confidence, Interest, and GEP (for the 10-item and 5-item scales). A Bonferonni-adjusted alpha value of .0125 was used to control for inflation of familywise type I error rates. No statistically significant differences were found between men and women on Confidence (p=.116, d=0.310), nor Interest, (p=.102, d=0.322). Though not statistically significant it is worth noting the effect sizes were of small to moderate size based on Cohen's (1988) benchmarks where men were nearly a third of a standard deviation higher than women in their confidence and interest in CS. Women were statistically significantly higher than men on GEP in CS for both the 10-item (p<.001, d=0.961) and 5-item (p<.001, d=0.837) factors, indicating women were more likely to believe women were equal to men in computing competence than men were by almost one standard deviation. This is considered to be a large effect size based on Cohen's (1988) benchmarks.

6 Discussion and future work

The purpose of the current study was to validate an attitudes measure for an undergraduate sample of non-CS majors. A secondary interest was to investigate gender differences on these constructs using the factor analysis results. Factor analysis results were consistent with those found by Heogh and Moskal in which a one-factor

Table 4: Descriptive and inferential statistics for gender comparisons on subscales

	women (n=49)		men (n=56)				
	Mean	Std. Dev.	Mean	Std. Dev.	t	p (two-tailed)	Cohen's d
Confidence	3.745	0.803	3.978	0.702	-1.58	0.116	0.310
Interest	3.404	0.839	3.656	0.746	-1.65	0.102	0.322
Gender Equality Perceptions (10-items)*	4.888	0.227	4.363	0.717	5.19	<.001	0.961
Gender Equality Perceptions (5-items)*	4.890	0.277	4.340	0.862	4.52	<.001	0.837

*Levene's test indicated the homogeneity of variances assumption was not met, thus Welch-Aspen t-test for unequal variances is reported.

solution best represented the constructs of Confidence, Interest, and what we refer to as Gender Equality Perceptions (GEP) in CS.

Of note, the current study provides support for a shortened 5-item GEP subscale indicating the GEP-5 performed as well as (and at times better than) the GEP-10. Results indicated the 5 positively worded items (GP3, GP4, GP8, GP9, and GP10) were sufficient to measure this construct and that the 5-item subscale was superior to the 10-item subscale in terms of the percent of variance explained in the variables (77% compared to 60%), the 5-item factor accounting for more variance in the items (i.e., higher communalities), and all items contributing to the overall reliability of the scale. Furthermore, both the 5-item and 10-item GEP subscales had Cronbach's coefficient alpha values of .919, also indicating the 5 negatively worded items may not be necessary to measure the construct. Also of note, the wording of the 5 negatively worded items is redundant with the 5 positively worded items in that the wording was just reversed for each item. Negatively worded items are often used in survey research to control for acquiescence. However, they often result in method effects where negatively worded items share some common variance above and beyond the factor, and thus can introduce additional measurement error, suggest an artificially inflated number of factors, and consequently reduce the validity of the interpretation drawn from scores. Based on these results, we recommend researchers use the 5 positively worded items from the GEP subscale.

As a secondary interest, results from the factor analysis were then used to examine gender differences for non-CS majors. women were nearly a full standard deviation higher on both the 10-item and 5-item GEP subscales (Cohen's d = .961 and .837, respectively). This is considered a large effect size in social science research and indicates women had higher ratings of gender equality in CS than men did. Though, not statistically significant, descriptively men were nearly a third of a standard deviation higher in Confidence and Interest than women (Cohen's d = .310 and .322, respectively).

As described at the beginning of this paper, there is a large body of work in literature related to gender gap in CS and lack of representation [4, 9, 34]. Previous studies have also consistently demonstrated that women tend to express lower confidence in their computing skills compared to men [14, 16]. Given the prior work on this topic, the result of our study is interesting as it indicates that female non-CS major students may not necessarily feel less confidence in their ability and they view women as capable of excelling in the CS field. This finding may be further evidence that the stereotype culture, lack of support and wavering sense of belonging is more pronounced for students majoring in CS, compared to those who are non-CS majors and only taking a programming course. Further analysis is needed to delve deeper into these issues and explore and compare the perceptions of students.

Findings from this study used convenience sampling, and as such our results might not be generalizable to all undergraduate students. However, the characteristics of our students are representative of typical undergraduate students majoring in CS at large research institutions in the United States. Furthermore, broadening the survey audience to include CS majors could offer valuable insights. This expansion is particularly relevant considering the limited variations observed in Non-CS majors. Including CS majors in the study could illuminate potential differences in experiences and attitudes within this specific academic discipline.

Several potential future directions for this study could significantly enhance the depth and accuracy of understanding participants' experiences and attitudes. Future studies should utilized confirmatory factor analysis (CFA) to examine the factor structure and provide more validity evidence to these findings. Additionally, validity evidence could be enhanced by examining criterion validity evidence. Also, using these factors to predict future outcomes for students may be helpful in providing validity evidence as well. Another potentially promising avenue is the incorporation of qualitative data, achieved through interviews with students. This approach aims to capture nuanced insights that may not be fully elucidated by quantitative measures alone. For example, conducting interviews with individuals from these groups could offer a richer understanding of this discrepancy.

7 Conclusion

In summary, the current study provided strong validity evidence for 1-factor subscales of Confidence in CS, Interest in CS, and Gender Equality Perceptions in CS. These subscales were found to have strong factor support and high internal consistency. The researchers suggest a shortened 5-item Gender Equality Perceptions in CS may be preferable to Hoegh and Moskal's (2009) original 10-item scale. Additionally, women were found to have much higher ratings (nearly a full standard deviation) of gender equality in CS than men.

References

[1] Deborah Bandalos and Sara Finney. 2019. *Factor Analysis: Exploratory and Confirmatory.* Publisher. Page Range pages. https://doi.org/10.4324/9781315755649-8
[2] Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.
[3] Maurice S Bartlett. 1951. The effect of standardization on a χ 2 approximation in factor analysis. *Biometrika* 38, 3/4 (1951), 337–344.
[4] Heather Blackburn. 2017. The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017. *Science & Technology Libraries* 36, 3 (2017),

235–273. https://doi.org/10.1080/0194262X.2017.1371658

[5] Urie Bronfenbrenner. 1979. *The ecology of human development: Experiments by nature and design.* Harvard university press.

[6] Patrick R. Brown and Holly M. Matusovich. 2013. Unlocking Student Motivation: Development of an Engineering Motivation Survey. In *Proceedings of the 120th Annual Conference and Exposition of the American Society for Engineering Education.*

[7] Tor Busch. 1995. Gender differences in self-efficacy and attitudes toward computers. *Journal of educational computing research* 12, 2 (1995), 147–158.

[8] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M Steele. 2009. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology* 97, 6 (2009), 1045.

[9] Sapna Cheryan, Sydney A. Ziegler, Angela K. Montoya, and Lily Jiang. 2017. Why are some STEM fields more gender balanced than others? *Psychological Bulletin* 143, 1 (2017), 1.

[10] Jacob Cohen. 1988. *Statistical power analysis for the behavioral sciences.* Routledge.

[11] Anna B Costello and Jason Osborne. 2019. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation* 10, 1 (2019), 7.

[12] A. B. Diekman and A. H. Eagly. 2000. Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Personality and Social Psychology Bulletin* 26, 10 (2000), 1171–1188. https://doi.org/10.1177/0146167200262001

[13] Brandeis Dorn and Allison E. Tew. 2015. Empirical Validation and Application of the Computing Attitudes Survey. *Computer Science Education* 25, 1 (2015), 1–36.

[14] Rodrigo Duran, Lassi Haaranen, and Arto Hellas. 2020. Gender differences in introductory programming: comparing MOOCs and local courses. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education.* 692–698.

[15] Allison Elliott Tew, Brian Dorn, and Oliver Schneider. 2012. Toward a validated computing attitudes survey. In *Proceedings of the ninth annual international conference on International computing education research.* 135–142.

[16] Katrina Falkner, Claudia Szabo, Dee Michell, Anna Szorenyi, and Shantel Thyer. 2015. Gender gap in academia: perceptions of female computer science academics. In *Proceedings of the 2015 ACM conference on innovation and technology in computer science education.* 111–116.

[17] Elizabeth Fennema and Julia A Sherman. 1976. Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for research in Mathematics Education* 7, 5 (1976), 324–326.

[18] Ashley V. Forssen, Barbara M. Moskal, and A. Robin Harriger. 2011. Measuring the Impact of a High School Intervention on Students' Attitudes in Information Technology: Validation and Use of an Attitude Survey. In *Proceedings of the American Society for Engineering Education.*

[19] Michail N. Giannakos, Ilias O. Pappas, Letizia Jaccheri, and Demetrios G. Sampson. 2017. Understanding Student Retention in Computer Science Education: The Role of Environment, Gains, Barriers and Usefulness. *Education and Information Technologies* 22, 5 (2017), 2365–2382. https://doi.org/10.1007/s10639-016-9538-1

[20] C. Good, A. Rattan, and C. S. Dweck. 2012. Why do women opt out? Sense of belonging and women's representation in mathematics. *J Pers Soc Psychol* 102, 4 (2012), 700–17. https://doi.org/10.1037/a0026659

[21] James M. Graham and Roberto Caso. 2002. Measuring Engineering Freshman Attitudes and Perceptions of Their First-Year Academic Experience: The Continuing Development of Two Assessment Instruments. In *Proceedings of the 32nd Annual Frontiers in Education Conference.*

[22] Daniel Heersink and Barbara M Moskal. 2010. Measuring high school students' attitudes toward computing. In *Proceedings of the 41st ACM technical symposium on Computer science education.* 446–450.

[23] Jonathan C. Hilpert, Glenda Stump, and Jenefer Husman. 2010. Pittsburgh Engineering Attitudes Scale - Revised: Evidence for an Improved Instrument. In *Proceedings of the 40th ASEE/IEEE Frontiers in Education Conference.*

[24] Larry S. Hirsch, John D. Carpinelli, Hersh Kimmel, Robert Rockland, and Jon Bloom. 2007. The Differential Effects of Pre-Engineering Curricula on Middle School Students' Attitudes to and Knowledge of Engineering Careers. In *Proceedings of the 37th ASEE/IEEE Frontiers in Education Conference.*

[25] Larry S. Hirsch, Scott J. Gibbons, Hersh Kimmel, Robert Rockland, and Jon Bloom. 2003. High School Students' Attitudes to and Knowledge About Engineering. In *Proceedings of the 33rd Annual Conference for Frontiers in Education.*

[26] Andrew Hoegh and Barbara M. Moskal. 2009. Examining science and engineering students' attitudes toward computer science. In *2009 39th IEEE Frontiers in Education Conference.* 1–6. https://doi.org/10.1109/FIE.2009.5350836

[27] Emma Hogan, Ruoxuan Li, and Adalbert Gerald Soosai Raj. 2023. Cs0 vs. cs1: Understanding fears and confidence amongst non-majors in introductory cs courses. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1.* 25–31.

[28] John L Horn. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30 (1965), 179–185.

[29] Lilly Irani. 2004. Understanding gender and confidence in CS course culture. In *Proceedings of the 35th SIGCSE technical symposium on Computer science education.* 195–199.

[30] Henry F Kaiser. 1974. An index of factorial simplicity. *psychometrika* 39, 1 (1974), 31–36.

[31] Carole L Kimberlin and Almut G Winterstein. 2008. Validity and reliability of measurement instruments used in research. *American journal of health-system pharmacy* 65, 23 (2008), 2276–2284.

[32] Özgen Korkmaz, Recep Çakir, and M Yaşar Özden. 2017. A validity and reliability study of the computational thinking scales (CTS). *Computers in human behavior* 72 (2017), 558–569.

[33] Sarah A. Lathem, Mark D. Neumann, and Nancy Hayden. 2011. The Socially Responsible Engineer: Assessing Student Attitudes of Roles and Responsibilities. *Journal of Engineering Education* 100, 3 (2011), 444–474.

[34] Kathleen J Lehman, Linda J Sax, and Hilary B Zimmerman. 2016. Women planning to major in computer science: Who are they and what makes them unique? *Computer Science Education* 26, 4 (2016), 277–298.

[35] Barbara M Moskal, Nathan Behrens, Mark Guzdial, Allison Elliott Tew, Wanda Dann, and Steven Cooper. 2006. Computer science assessment instrument development: Evaluating attitudes and outcomes. In *STEM Assessment Conference.* 194.

[36] Jean Nocito-Gobel, Michael A. Collura, Susan Daniels, and Ibrahim I. Orabi. 2005. Are Attitudes Toward Engineering Influenced by a Project-Based Introductory Course?. In *Proceedings of the 2005 American Society for Engineering Education Conference and Exposition.* Portland, OR.

[37] Brian P O'connor. 2000. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior research methods, instruments, & computers* 32, 3 (2000), 396–402.

[38] Arif Rachmatullah, Eric Wiebe, Danielle Boulden, Bradford Mott, Kristy Boyer, and James Lester. 2020. Development and validation of the Computer Science Attitudes Scale for middle school students (MG-CS attitudes). *Computers in Human Behavior Reports* 2 (2020), 100018.

[39] Vennila Ramalingam and Susan Wiedenbeck. 1998. Development and validation of scores on a computer programming self-efficacy scale and group analyses of novice programmer self-efficacy. *Journal of Educational Computing Research* 19, 4 (1998), 367–381.

[40] C. Riegle-Crumb, B. King, and Y. Irizarry. 2019. Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields. *Educational Researcher* 48, 3 (2019), 133–144. https://doi.org/10.3102/0013189X19831006

[41] Katharine Semsar, Jennifer K Knight, Gülnur Birol, and Michelle K Smith. 2011. The Colorado learning attitudes about science survey (CLASS) for use in biology. *CBE—life sciences education* 10, 3 (2011), 268–278.

[42] Jane Sinclair and Sara Kalvala. 2015. Exploring societal factors affecting the experience and engagement of first year female computer science undergraduates. In *Proceedings of the 15th Koli Calling Conference on Computing Education Research.* 107–116.

[43] Nicole M Stephens, Hazel Rose Markus, and Stephanie A Fryberg. 2012. Social class disparities in health and education: reducing inequality by applying a sociocultural self model of behavior. *Psychological review* 119, 4 (2012), 723.

[44] Una Tellhed, Martin Bäckström, and Fredrik Björklund. 2018. The Role of Ability Beliefs and Agentic vs. Communal Career Goals in Adolescents' First Educational Choice. What Explains the Degree of Gender-Balance? *Journal of Vocational Behavior* 104 (2018), 1–13. https://doi.org/10.1016/J.JVB.2017.09.008

[45] Meng-Jung Tsai, Ching-Yeh Wang, and Po-Fen Hsu. 2019. Developing the computer programming self-efficacy scale for computer literacy education. *Journal of Educational Computing Research* 56, 8 (2019), 1345–1360.

[46] Wayne F. Velicer. 1976. Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika* 41 (1976), 321–327.

[47] Jennifer Wang, Hai Hong, Jason Ravitz, and Marielena Ivory. 2015. Gender differences in factors influencing pursuit of computer science and related fields. In *Proceedings of the 2015 ACM conference on innovation and technology in computer science education.* 117–122.

[48] Dana Wanzer, Tom McKlin, Doug Edwards, Jason Freeman, and Brian Magerko. 2019. Assessing the attitudes towards computing scale: A survey validation study. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education.* 859–865.

[49] Eric Wiebe, Laurie Ann Williams, Kai Yang, and Carol S Miller. 2003. *Computer science attitude survey.* Technical Report. North Carolina State University. Dept. of Computer Science.

[50] Brenda Cantwell Wilson. 2002. A study of factors promoting success in computer science including gender differences. *Computer Science Education* 12, 1-2 (2002), 141–164.

[51] Angela Zavaleta Bernuy and Brian Harrington. 2020. What are we asking our students? A literature map of student surveys in computer science education. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education.* 418–424.