ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #37074

Toward Diversifying Computer Science With Novel Interest-Based Models of Students

Joshua Gross

Joshua Gross is an assistant professor of computer science at CSUMB. He spent nearly a decade as a software engineer, earning an MS in software engineering from the University of St. Thomas in St. Paul, Minnesota. He holds a PhD in information sciences and technology from Penn State, where his research focused on the intersection of artificial intelligence and human-computer interaction. His current research is focused on the psychology of programming, with the goal of better understanding factors that support student success in undergraduate computer science coursework and in software development professions, with the hope of addressing the severe shortfall of qualified developers.

Keyoni McNair

Keyoni McNair is an incoming Software Developer at SAP. She recently completed a Bachelor of Science degree in Computer Science with a Concentration in Software Engineering at California State University, Monterey Bay. She was a founding officer in an ACM-W Chapter and a teaching assistant for two upper-division computer science courses. Her research interests include diversity in computer science, retention in lower division computer science courses, and promotion of computer science education.

Toward Diversifying Computer Science With Novel Interest-Based Models of Students

Abstract

Computing disciplines face a profound diversity challenge, and recent efforts have not fundamentally altered the demographics of the field. Part of the challenge comes from stereotypes of computer science students that emphasize a narrow set of interests that is demonstrably alienating to many of the students we wish to recruit and retain. This paper describes the challenges that stereotypes represent, and describes a novel approach to understanding computer science students based on shared interests in diverse areas. The development of the survey instrument is described, along with initial tests of validity and descriptive statistics reported from students majoring in computer science and other fields.

Introduction

Computing disciplines, as with many engineering disciplines, face a profound diversity challenge: white and Asian men are over-represented among students, professionals, and faculty, with other groups underrepresented. Computing faces the additional challenge of a decade of surging demand for professionals and increased enrollment in academic programs. Data from the ACM NDC [1] and Taulbee [2] surveys shows improved representation in awarded bachelor's degrees, but the shift is not dramatic.

One aspect of the larger problem stems from the fact that computer science (CS) students are often perceived through the lens of stereotypes. Research has shown that CS majors are seen as mostly nerdy, lazy, introverted white and Asian men with a strong interest in mathematics [3, 4, 5]. These stereotypes are common among both CS students and the larger world, and undermine efforts to diversify student populations because they discourage some students from entering the field who might otherwise be interested in CS. In addition, these stereotypes feed into identity concerns such as stereotype threat and imposter phenomenon [6], which can interfere with a student's resilience.

These stereotypes are problematic, but we face an additional challenge: the statistics affirm the demographic stereotypes. In a preliminary study of 11 factors that influence CS student decisions to continue in the major, being male was (by far) the most significant factor [7]. While the long-term goal is for CS students to feel represented, that is often unrealistic in the short term. One of the authors of this paper lamented being the only black student in her course on race and gender in technology. We struggle with a chicken-and-egg problem: potential students feel alienated

by the lack of representation, and because they then choose other fields, we cannot address the representation issue. We must become more diverse, but that is not a directly actionable goal. We cannot simply materialize a more diverse student population; we must improve recruitment and retention.

While demographics are important to identity, interests are part of identity formation, too. By considering CS students' interests in a broad range of activities, we can help potential students see that students in CS share some interests with them, even if they struggle to find peers and mentors who represent their race, gender, disability status, or other factors relevant to them. Exploring this will allow us to show diversity of interests, one of the few areas in which CS is diverse.

Engineering is defined by the students who become engineers, and by the people who educate those students. The question of "What is engineering?" cannot be answered without also answering "Who are engineers?" Our hope is that we can improve the discipline by providing a more robust answer to the latter question, and perhaps address demographic under-representation.

Our models and understanding of CS students are relatively threadbare. Research on students in STEM disciplines often focuses on issues of academic preparedness and ability. Analysis of characteristics that are predictive of success tend to focus on obvious qualities like prior programming experience or interest in areas obviously and directly related to the field, such as mathematics. This narrow set of dimensions is reductive and casts a narrow net for recruiting efforts.

This paper describes an exploratory approach that evaluates student interest in several unconventional areas (in addition to mathematics), built partially on existing work and extended to better understand CS students relative to the stereotypes. The goal of this research is to better understand the common threads of interest among CS students beyond the stereotypes of mathematics, video games, and other computer-related factors. By more broadly exploring the interests of CS students, we hope this paper can be one step to producing better models and understanding of our students. Further, we hope that such a model will allow us to better recruit and retain students who do not see themselves as belonging in computer science, but who would have the potential to succeed. We believe that such an outcome would increase the presence and success of students from underrepresented groups.

To begin building a better model of student interest, we developed an interest inventory addressing five different interest areas (mathematics, writing, creative hobbies, athletics, and public performance). Our goal is to offer a more complete understanding of our students, to see them in light of their differences in addition to their similarities. We surveyed undergraduate students (both CS majors and non-CS majors) at a regional comprehensive university. This paper presents the origins and design of our instrument, the initial data collection efforts, current analysis, and plans for future work.

Literature Review - The Challenge of Stereotypes

We begin by exploring existing attempts to understand how CS students see themselves and their peers, especially in light of stereotypes held by people in and outside of the field. Stereotypes can undermine the development of a sense of belonging in the field, leading to low persistence. Stereotypes can also alter how people are treated, but the focus of this work is on self-perception

and our understanding of student identity, rather than on biased behavior.

Past studies have shown that potential and current CS students hold stereotypes of CS students as asocial, competitive, "geeky" or "nerdy" white or Asian men who love mathematics and programming [3, 8, 4, 5]. Clearly, this is not representative of everyone in the field, and these stereotypes can hinder diversity efforts by undermining a student's perception of belonging and thus their likely retention in the field.

The presence and consequences of gender stereotypes specifically have been studied extensively. Kapoor and Gardner-McCune demonstrated that "Presence of gender bias in the classroom" was one of the most significant factors that influenced students' decision to switch their major away from computer science [9]. In 2009, Cheryan *et al.* conducted an experiment in which some students were exposed to classrooms designed along gender and "geek" stereotypes, including posters of sports and science fiction themes, while others were exposed to classrooms without these stereotypes represented; they found that the stereotype-laden rooms led to women expressing less identification with (and interest in) CS, while having no effect on men [10].

In a later experiment, Cheryan *et al.* exposed students to one of two role models: either one matching a CS stereotype (hobbies of playing video games and programming) or one not matching these stereotypes (hobbies of playing sports and socializing). They found that women faced with a stereotypical model reported a reduced interest in CS and a lower sense of belonging compared to a baseline of women not exposed to the role model[11]. Winter *et al.* conducted in-depth interviews with first year women CS students. Many of these women cited knowledge of the "male and geeky" stereotypes, gendered language in classrooms, and instances of feeling discriminated against [12].

This sense of belonging is important. Veilleux *et al.* found that a student's *perception* of their ability strongly correlated with their sense of belonging, even more so than their grades. Students who remained felt supported, accepted, and comfortable in their class and major [13].

Imposter phenomenon (IP), also known as imposter syndrome, describes an intense feeling of not belonging and fraudulent presence. While research demonstrates that women in CS do experience IP more than men, both groups experience high levels [6], and CS students have higher levels of IP compared to other fields [14]. IP can lower a student's ability to demonstrate resilience, a core metacognitive method for succeeding in the face of challenge and critique.

It should be noted that tereotypes are complex phenomena, and can have negative and positive impacts in different situations and for different groups and sub-groups of students. Winter *et al.* saw the hedging women did to minimize their experiences with stereotypes as connected to their efforts to demonstrate resilience and persist in the major [12]. In Cheryan's 2009 experiment, sports were classified as part of the CS stereotype, while in the later experiment, sports were used to contrast the stereotype [10, 11]. Similarly, competitiveness is seen as a stereotype that creates problems for women in the research conducted by Lewis *et al.* [4], but it is important to not over-generalize stereotypes, such as assuming that all groups of women are more interested in cooperation than competition. Stewart *et al.* [15] found that women student athletes in CS academically outperformed non-athlete women and all men studying CS, and that these women evidenced a higher level of competitiveness than non-athlete women CS students.

Through these studies it is clear that stereotypes can cause adverse reactions to CS from potential students and hinder a sense of belonging for current students. This creates a distinct barrier between potential students, current CS students, and their connection to CS. People who are interested in CS are not a monolith, and it is easy to see from this literature how ineffective a monolithic model is in increasing the diversity in the field. Even so, categorizing certain interests as stereotypical (and thus concerning) is also problematic, since students from underrepresented groups are also not monolithic and may share these interests.

These studies have mostly demonstrated the power of stereotypes to deter a more diverse population from pursuing and persisting in CS. What we lack is an alternative model focused on attracting students and helping improve representation.

A New Approach to Modeling Students

Our goal is to build a new set of models of CS students, focused on shared interests; we believe that these new models (and supporting data) will enhance efforts to bring in a more diverse and representative student body. That goal is challenging in the face of current demographics, because of *homophily*, the human tendency to associate with people similar to us. One form of similarity is membership in groups based on race, ethnicity, gender, sexual orientation, and many other factors. However, interests are another key component of identity and can serve as a foundation for homophily. Our shared interests allow us to connect with peers and mentors, even if those peers and mentors do not share many aspects of identity.

Further, there is a connection between identity and resilience. Students who believe that they belong in CS are more likely to demonstrate resilience in the face of the nearly inevitable challenge that they will experience. This resilience can increase academic success.

One of the challenges in recruiting students to study CS is that most middle and high school students do not have access to CS classes due to the lack of sufficient teaching resources [16]. DeClue and Provance [17] explored related interests that could be identified in students as early as middle school to encourage them to pursue or explore CS. These three interest areas, mathematics, language, and creativity, serve as a foundation for our own research. They conducted an interview study with existing upper-division CS students to explore the students' relationship with each interest from when they were in middle school, rather than currently. All students expressed an interest in at least one of the areas; some expressed multiple such interests.

Typical positive responses for math included expressing love for the subject, talking about high grades, and understanding the subject even if it was not loved. For language, students responded with their love of reading and above-average amount of reading when compared to their peers. When talking about creativity, students spoke of different ways they tinkered, such as taking objects apart and reassembling them and learning maker hobbies, like cross-stitch or Photoshop.

The present work began to consider the breadth of interests that CS students have. While interest in mathematics falls into the classic group of CS stereotypes, it is relevant and can serve as a baseline. Language and creativity also have deep connections to CS, but these connections are often not visible to people outside of the field. Students interested in language and creative hobbies might not be encouraged to study CS, while students interested in mathematics are often encouraged to

do so.

Others have considered similar approaches in conducting research about broadening perceptions of engineering. For language, Halada and Khost engaged a diverse group of students in STEM concepts by having students share their narratives of their own personal use of technology and reading and writing science fiction [18]. For creative hobbies, Garlock *et al.* taught the concept of resonance to students by having them construct buildings out of spaghetti and marshmallows; this course was taken by students from STEM and non-STEM majors, and all students reported increased perception of engineering as a creative profession [19].

More radically, Flath (a mathematician with an engineering background) and Michelfelder (a philosopher interested in the philosophy of engineering and technology) developed a course entitled "Thinking Like an Engineer" for a liberal arts college with no engineering majors. The authors (who were also the instructors) focused this first-year course on helping students understand the importance of connecting technical and nontechnical skills and intelligence, combined with their knowledge of the world, to solving engineering problems [20]. They combined STEM and non-STEM instruction to connect engineering thinking to a broader set of intellectual skills and interests. Focusing solely on the ability to master technical material fails to address this and produces engineers with limits and blindspots, rather than the intellectually and socially diverse workforce necessary to solve 21st century engineering challenges. The course leveraged student interests beyond technical topics to better engage the students in technical problem solving.

Developing the Instrument

We chose to build off of DeClue and Provance's work, constructing an inventory to evaluate student interest in a variety of both obvious (e.g., mathematics) and non-obvious (e.g., writing) categories. In addition to their three categories (Mathematics, Language, Creativity), we added two categories, Athletics and Public Performance.

We had three concerns about asking students direct questions asking their interest in each of our five areas. First, we were concerned that students would feel pressured to identify with what they thought CS students should be interested in. Second, we felt that the labels themselves would lead to bias in the student experience of completing the inventory. Third, because activities are concrete and actionable, they may more clearly express the intent of the instrument and more authentically reflect a student's interest.

Using the US Department of Labor's MyNextMove [21] interest inventory as a model to develop the instrument, we developed our own instrument to evaluate student interest in the five categories. MyNextMove is a research-driven tool provided by the US Department of Labor to help people explore career options based on interests. Participants express their interest in a variety of job-related activities using a Likert-like visual scale. Examples of activities include "Repair household appliances", "Study the movement of planets", and "Compose or arrange music". Responses are scored by category, and these scores are used to match the respondent to possible careers based on prior data.

We iteratively developed five activities for each category, ranging from short duration (e.g., "Write a poem, short story, or article" or "Play catch, frisbee, or a similar game with friends") to a much

larger commitment (e.g., "Take a technical or creative writing class not your major" or "Join an intramural sports team"). In operationalizing the language interest category, we chose to focus on writing, since that is both active and particularly relevant to CS. We asked respondents to indicate how much they would enjoy each activity using a Likert-like scale, with options of Strongly Dislike, Dislike, Neither Like or Dislike, Like, Strongly Like. The questions were grouped by duration, rather than category, starting with shorter activities and ending with the longest activities. The full instrument is presented in the appendix.

Our belief was that, by offering common activities associated with the interests, we could better evaluate a student's actual interest. This was partly inspired by Bandura's model self-efficacy [22], which focuses on a person's self-perception of their ability to perform a concrete task (e.g., write a for loop), rather than their generic evaluation of their level of skill.

Collecting Data

In order to evaluate the quality of the instrument, several drafts of the core survey questions (the interest inventory itself) were completed by acquaintances. Overall, a group of 39 respondents (including college students and other adults not enrolled in college) completed various drafts of the survey. The focus of the iterative design and testing was to demonstrate consistency within each category and ensure that respondents saw the activities as representative of the categories.

The survey was conducted at primarily undergraduate Hispanic-serving institution in coastal rural California. After completing an IRB, we distributed the survey to 92 students in Spring 2021 who were enrolled in either a first- or second-semester (introductory or intermediate) CS course. Of that pool, we had responses from 69 students, of which 63 were usable (6 incomplete, 68% effective response rate). These respondents are labeled *CS* students.

In Spring 2022, we distributed the instrument to 130 students in a general education course on film history. We had responses from 40 of which 37 were usable (3 incomplete). Of these, ten identified as CS majors or minors, and we eliminated those from our analysis, leaving 27 responses (21% effective response rate). These are labeled *non-CS* students.

Select self-reported demographics of respondents are reported in Table 1. All demographic questions were optional, although most respondents answered all questions.

Evaluating the Instrument

For each item, we translated from the Likert-like scale to a numerical scale of -2 to +2, with -2 representing Strongly Dislike and +2 representing Strongly Like. We then summed the scores on a per-category basis. A respondent who strongly disliked all activities in a category would score -10, and a respondent who strongly liked all activities in a category would score +10. Each respondent had a separate score for each category. We call this *inferred interest*.

In order to gauge whether our groups of activities matched our subjects' perceptions of the intended categories we directly asked each respondent to indicate their interest in the category (also on a Likert scale) at the end of the survey. We called this the *self-reported interest*. Since the respondents completed this part of the survey after the interest inventory, we felt our concerns about

Table 1: Participant Demographics

	CS	Non-CS
Age		
18-24 years old	90%	93%
25-44 years old	10%	7%
Gender		
Male	69%	33%
Female	26%	56%
Non-binary	3%	11%
Race		
American Indian/Alaskan Native	0%	0%
Asian	14%	23%
Black or African American	8%	4%
Native Hawaiian or Pacific Islander	0%	0%
White	59%	54%
Two or more races	19%	15%
Other	0%	4%
Ethnicity		
Hispanic, Latinx, or Spanish Descent	50%	44%
Not Hispanic, Latinx, or Spanish Descent	50%	56%

Table 2: Cronbach's α for CS and Non-CS Students

	n	Math	Writing	Creativity	Athletics	Performance
CS	63	.72	.79	.63	.81	.70
Non-CS	27	.87	.73	-.19	.74	.56
Both	89	.79	.8	.51	.8	.68

bias would not materialize.

Our first analysis concerns validating the instrument. We chose two methods: a Cronbach's α test of the internal consistency of each category, and a Spearman's correlation coefficient between the inferred interest and the self-reported interest.

Using Cronbach's α, we looked at the reliability of the activities for each category for each group of respondents as seen on Table 2. While most of the categories fell into the acceptable range, the reliability of Performance with non-CS students is poor, and Creativity is dramatically off (possibly due to small sample size).

For the Spearman's correlation, the null hypothesis is that there would not be a statistically significant correlation between the inferred and self-reported interest in any of the respondent groups. As seen in Table 3, there are significant correlations between these two variables in every category except for Creativity with the non-CS students. This supports the claim that our model of inferred interest is consistent with genuine interest, yet left room for improvement for future iterations, particularly with creativity and public performance.

Table 3: Correlation Between Inferred and Self-Reported Interests

	CS (df= 62)		Non-CS (df=25)		Both (df=88)	
	Correlation Coefficient	p-value	Correlation Coefficient	p-value	Correlation Coefficient	p-value
Math	.63	$< .0001$	.84	$< .0001$	.65	$< .0001$
Writing	.82	$< .0001$	.75	$< .0001$	.76	$< .0001$
Creativity	.56	$< .0001$	.35	$< .0792$	.55	$< .0001$
Athletics	.74	$< .0001$	.69	$< .0001$	.66	$< .0001$
Performance	.52	$< .0001$	.64	$< .0004$	.53	$< .0001$

Note the significantly higher correlation for mathematics among the non-CS students. As will be demonstrated below, CS students had dramatically higher self-reported interest in math, compared to their inferred interest. This tends to support our concern about bias encouraging CS students to self-identify as interested in mathematics, despite little interest in mathematics activities.

Analysis of Data

Having satisfied our concerns that our instrument offered meaningful insight, we next analyzed inferred and self-reported interest in each category. DeClue and Provance's original research suggests that mathematics, creativity, and writing will all score highly, and the common stereotypes of CS students would predict low scores for Athletics and Public Performance. However, since our goal was to establish a baseline understanding of student interest, we did not treat these as research hypotheses.

We explored descriptive statistics for each category for CS and non-CS students in Table 4 and Figure 1. CS students had Mathematics and Writing inferred interest scores near the middle of the scale, with similar standard deviations, indicating that respondents tended to be neutral toward those two areas, although this may reflect that they are *academic* areas, rather than hobbies. Creativity had the highest average score among the CS respondents, followed by Athletics coming up afterwards. Athletics was a surprising result as it was one of the areas we assumed not to be connected to CS students. Public Performance, however, met our expectations as it has the lowest score in the group.

The means for self-reported and inferred interest were generally similar for both groups with three notable exceptions. First, CS students self-reported significantly higher interest in Mathematics than they expressed interest in Mathematics-oriented activities. Second, both groups expressed significantly higher self-reported interest in Public Performance than their expressed interest in the associated activities.

The most dramatic differences between the two groups of students in in Mathematics and Writing. It is unsurprising that non-CS students have more strong negative feelings about mathematics. The much stronger interest in Writing seemingly reflects the gap that CP Snow explored in his essay, "The two cultures" [23], but given the importance of writing to CS, this may be a lever with which to engage students who might not otherwise consider CS.

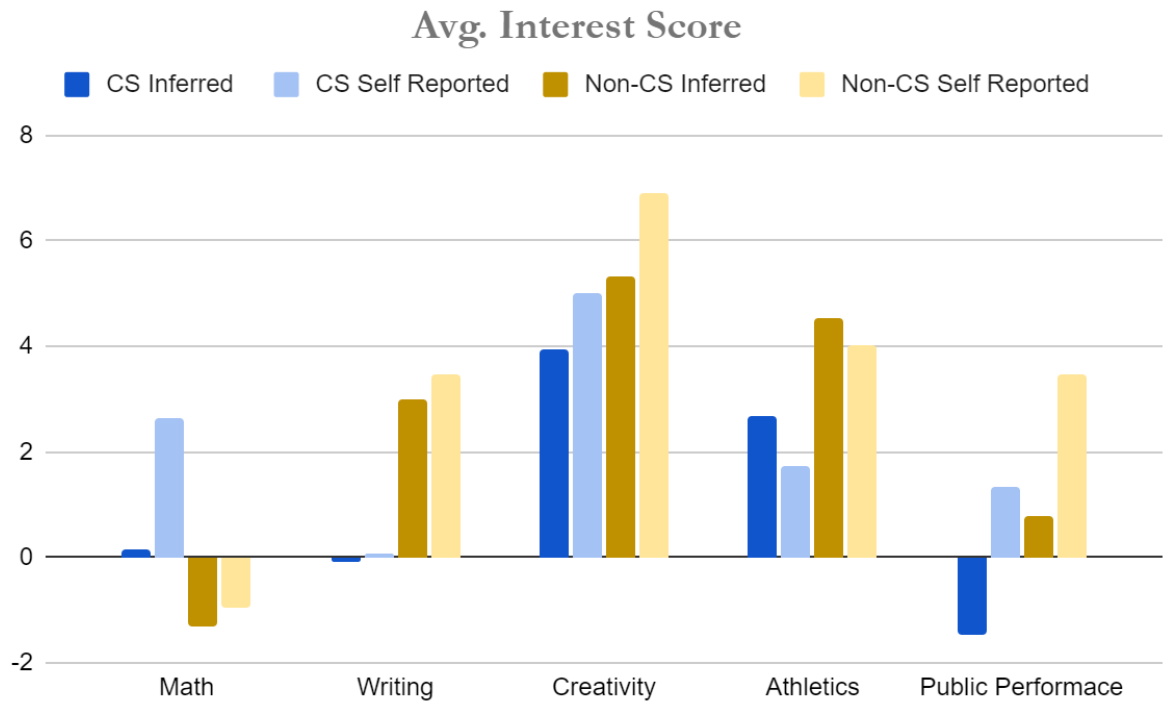

Figure 1: Average Inferred and Self-Reported Interest Scores

There are several possible interpretations of the deviation in Mathematics. We will need to investigate further to better understand the issue. For understanding our students, insight will be important, but for engaging potential students through interest-related homophily, projected interest may be more important.

- The specific activities in the survey were not of interest to the CS students
- The CS students are interested in mathematics in the abstract, but are generally not interested in mathematics activities
- Our concern about students feeling compulsion to identify as interested in mathematics was warranted (this may also explain the higher inferred interest for Athletics, but the difference is less dramatic)

It is possible that the Public Performance activities are not representative of how the participants see that category, and as discussed above, we intend to thoroughly review Creativity and Public Performance categories.

Limitations and Future Work

Our study has some significant limitations due to the incipient nature of the instrument and limits of our data collection to date. In addition, more analysis of our existing data may yet yield more insights.

Table 4: Average Inferred and Self-Reported Interest Scores for CS and Non-CS Students

	CS Students				Non-CS Students			
	Mean (*SD*)				Mean (*SD*)			
Category	Inferred		Self-Reported		Inferred		Self-Reported	
Math	0.17	*(3.2)*	2.62	*(4.6)*	-1.30	*(4.6)*	-0.96	*(6.8)*
Writing	-0.09	*(3.9)*	0.08	*(5.2)*	3.00	*(4.1)*	3.46	*(6.3)*
Creativity	3.94	*(2.7)*	5.00	*(4.8)*	5.33	*(2.1)*	6.92	*(4.0)*
Athletics	2.67	*(3.7)*	1.75	*(5.4)*	4.52	*(3.6)*	4.04	*(5.5)*
Performance	-1.47	*(3.8)*	1.35	*(5.5)*	0.78	*(4.3)*	3.46	*(6.0)*

We have reflected on a number of limitations of our instrument:

- Two categories (Mathematics and Writing) can be considered academic, while the others can be considered as hobbies
- We developed a category for puzzles, games, and problem solving, but were not able to produce consistency, and we hope to incorporate that in subsequent revisions
- Some categories demonstrated weaker reliability and correlation to self-reported interest
- Some large gaps between self-reported and inferred interest remain unexplained
- We would like to add more academic and non-academic interests

In addition, all of our data were collected from students taking an introductory or intermediate course in the same semester (to avoid multiple entries from a single student), and all non-CS students were solicited from one general education course. Our data may be characteristic of our student population; for example, the high interest in athletics may be a reflection of why some students chose this an institution located in an area with many options for outdoor physical activity (or perhaps it is merely a function of age). We hope to gather more data from students in CS and other majors, across years, and from multiple institutions. We also hope to expand to students in other engineering majors; no such majors are offered at our institution, but as we expand to other institutions, this will be possible.

Beyond descriptive statistics on a per-category basis, we hope that further analysis will help us identify connections between interests. Perhaps CS students can be described in terms of different subgroups with a set of shared positive and negative interests, and we intend to use cluster analysis to derive such possible groups. For example, we may see a cluster of students who are interested in writing, but are not interested in math. We hope that this analysis will allow for more three-dimensional models of the types of students who pursue CS.

Conclusions

As noted in the introduction, our efforts to understand our students have traditionally focused on topics we believe to be directly relevant to our field. Mathematics and prior programming experience may be predictive of interest and success, but it seems likely that those students are already considering CS.

We believe that this research has the potential to show our students in a new light, and to allow us to show potential students that CS majors share interests with them. While our results are preliminary, some of the results are surprising. CS students may be more ambivalent toward math than the stereotypes indicate, and more interested in athletics. Our small sample of non-CS students was far more positive about writing than they were negative about mathematics. These data are far from conclusive, but they demonstrate some interesting questions to explore. Further, we believe that this research has the potential to provide new insight into the nature of engineering through a better understanding of engineers.

Of course, the primary purpose of the current work is to explore the need for new models of CS, and a novel approach to build those models based on their shared interests. It is our hope that further work will provide more insight and opportunities to use our knowledge to build a more diverse and representative future.

References

[1] S. Zweben, J. Tims, and Y. Timanovsky, "ACM-NDC study 2019—2020: eighth annual study of non-doctoral-granting departments in computing," *ACM Inroads*, vol. 11, no. 3, pp. 26–37, Aug. 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3410475

[2] S. Zweben and B. Bizot, "2017 CRA Taulbee survey," *Computing Research News*, vol. 30, no. 5, pp. 1–47, 2018. [Online]. Available: https://cra.org/resources/taulbee-survey/

[3] E. M. Mercier, B. Barron, and K. M. O'connor, "Images of self and others as computer users: The role of gender and experience," *Journal of Computer Assisted Learning*, vol. 22, no. 5, pp. 335–348, 2006, publisher: Wiley Online Library.

[4] C. M. Lewis, R. E. Anderson, and K. Yasuhara, "'I Don't Code All Day': Fitting in Computer Science When the Stereotypes Don't Fit," in *ICER '16*. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 23–32. [Online]. Available: https://doi.org/10.1145/2960310.2960332

[5] A. Alshahrani, I. Ross, and M. I. Wood, "Using social cognitive career theory to understand why students choose to study computer science," in *ITiCSE '18*, 2018, pp. 205–214.

[6] A. Zavaleta Bernuy, A. Ly, B. Harrington, M. Liut, A. Petersen, S. Sharmin, and L. Zhang, "Additional Evidence for the Prevalence of the Impostor Phenomenon in Computing," in *SIGCSE '22*, Providence, RI, 2022.

[7] L. J. Barker, C. McDowell, and K. Kalahar, "Exploring factors that influence computer science introductory course students to persist in the major," *ACM SIGCSE Bulletin*, vol. 41, no. 1, pp. 153–157, Mar. 2009. [Online]. Available: https://doi.org/10.1145/1539024.1508923

[8] S. Cheryan, A. Master, and A. N. Meltzoff, "Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes," *Front. Psychol.*, vol. 6, p. 49, Feb. 2015, publisher: frontiersin.org. [Online]. Available: http://dx.doi.org/10.3389/fpsyg.2015.00049

[9] A. Kapoor and C. Gardner-McCune, "Considerations for switching: exploring factors behind CS students' desire to leave a CS major," in *ITiCSE '18*. New York, NY, USA: Association for Computing Machinery, Jul. 2018, pp. 290–295. [Online]. Available: https://doi.org/10.1145/3197091.3197113

[10] S. Cheryan, V. C. Plaut, P. G. Davies, and C. M. Steele, "Ambient belonging: how stereotypical cues impact gender participation in computer science," *Journal of personality and social psychology*, vol. 97, no. 6, p. 1045, 2009, publisher: American Psychological Association.

[11] S. Cheryan, B. J. Drury, and M. Vichayapai, "Enduring influence of stereotypical computer science role models on women's academic aspirations," *Psychology of women quarterly*, vol. 37, no. 1, pp. 72–79, 2013, publisher: Sage Publications Sage CA: Los Angeles, CA.

[12] E. Winter, L. Thomas, and L. Blair, "'It's a Bit Weird, but it's OK'? How Female Computer Science Students Navigate being a Minority," in *ITiCSE '21*. New York, NY, USA: Association for Computing Machinery, Jun. 2021, pp. 436–442. [Online]. Available: https://doi.org/10.1145/3430665.3456329

[13] N. Veilleux, R. Bates, C. Allendoerfer, D. Jones, J. Crawford, and T. Floyd Smith, "The relationship between belonging and ability in computer science," in *SIGCSE '13*, 2013, pp. 65–70.

[14] S. Rosenthal and T. R. Chung, "A Data Science Major: Building Skills and Confidence," in *SIGCSE '20*. New York, NY, USA: Association for Computing Machinery, Feb. 2020, pp. 178–184. [Online]. Available: https://doi.org/10.1145/3328778.3366791

[15] K. J. Stewart, G. C. Townsend, and S. Tunguz, "An Unexplored Relationship: Women in Computing and Athletics," in *SIGCSE '20*. New York, NY, USA: Association for Computing Machinery, Feb. 2020, pp. 699–705. [Online]. Available: https://doi.org/10.1145/3328778.3366884

[16] Google and Gallup, "Diversity gaps in computer science: exploring the underrepresentation of girls, Blacks and Hispanics," 2016. [Online]. Available: https://www.voced.edu.au/content/ngv:76007

[17] T. DeClue and M. Provance, "Predicting digital success: identifying computer science talent in K12 students," *Journal of Computing Sciences in Colleges*, vol. 33, no. 5, pp. 108–114, 2018.

[18] G. P. Halada and P. H. Khost, "The Use of Narrative in Undergraduate Engineering Education," in *ASEE 2017*, 2017.

[19] M. E. Garlock, A. Bhatia, and E. H. Laffey, "Effective approaches for teaching STEM-literacy for all majors: The example of resonance," in *ASEE '17*, 2017.

[20] D. Flath and D. P. Michelfelder, "Thinking Like an Engineer: Responding to the Engineering Blind Spot at Liberal Arts Colleges," in *ASEE '17*, 2017.

[21] "O-NET Interest Profiler at My Next Move." [Online]. Available: https://www.mynextmove.org/explore/ip

[22] A. Bandura, W. H. Freeman, and R. Lightsey, "Self-Efficacy: The Exercise of Control," *Journal of Cognitive Psychotherapy*, vol. 13, no. 2, pp. 158–166, Jan. 1999. [Online]. Available: https://connect.springerpub.com/content

[23] C. P. Snow, *The two cultures*. Cambridge University Press, 2012.

Appendix

Activities

Participants were given a list of activities from each topic and a 5-point Likert-like scale (Strongly Dislike, Dislike, Neither Like or Dislike, Like, Strongly Like). For clarity, activities are grouped according to topic, not by duration as presented to participants.

How much would you enjoy these activities?

- Math
  - Manually calculate a tip at a restaurant
  - Help a friend with a math homework problem
  - Study more in a math class to get a better grade
  - Join a math club
  - Take an additional math class outside of major requirements
- Language / Literature / Writing

  - Write a poem, short story, or article
  - Give a friend feedback on an essay
  - Submit an early draft of a writing assignment to get additional feedback
  - Join a book or writing club
  - Take a technical or creative writing class not your major
- Creativity/ Makers
  - Cook a dish from a recipe
  - Build a Lego set (or similar building toys)
  - Create an object by hand (e.g., woodworking or knitting)
  - Compose a song
  - Learn a new hobby (e.g., learn a new instrument or type of art)
- Sports / Athletics
  - Play catch, frisbee, or a similar game with friends
  - Go for a run, walk or bike ride
  - Participate in a 5k or fun run
  - Practice extensively to improve your performance in your physical activity or sport of choice
  - Join an intramural sports team
- Performance
  - Dance socially
  - Give a presentation about a topic of your choice to a group of interested students
  - Attend an improv workshop
  - Join a debate team
  - Perform a play on stage
- Puzzle and Games/ Problem Solving *(Removed in final iteration)*
  - Try to solve a crossword puzzle, word search or sudoku
  - Try to solve an escape room
  - Play a strategy board game (e.g., Chess)
  - Play a video game for an hour
  - Diagnose and fix a mechanical or computer problem

How would you rank your interest in activities in these categories? Participants were given a 5 point scale ranging from 'Strongly Dislike' to 'Strongly Like' to rank their interest in 'Math', 'Writing', 'Creativity', 'Sports/Athletics', 'Performance'

Are you currently a Computer Science Major or Minor? (Yes/No)

How much do you agree with the following statements? Participants were given a 5 point scale ranging from Strongly Disagree to Strongly Agree

- Computer science is interesting to me
- I plan to study (or continue to study) in computer science