Who Are You? We Really Wanna Know... Especially If You Think You're Like a Computer Scientist

Rob Semmens
Graduate School of Education
Stanford University
semmens@stanford.edu

Chris Piech
Dept of Computer Science
Stanford University
piech@cs.stanford.edu

Michelle Friend
Graduate School of Education
Stanford University
mfriend@stanford.edu

ABSTRACT
We developed a short, easily implemented survey that measures the similarity in phrases describing the self and a computer scientist. Additionally, we took initial steps in determining adjectives or phrases that describe a stereotypical computer scientist. We then administered this survey before and after an eight-week summer computer science program for high school girls. We found that phrases or adjectives used to describe the self converged with those to describe the computer scientist. In addition, descriptions of both were more positive at the end of the program compared to the beginning. Finally, the stereotypical of a computer scientist decreased from the beginning to the end of the program. Future work includes refinement of the stereotype measure and assessing different types of computer science programs.

Categories and Subject Descriptors
K.3.2 [**Computer science education**]: Metrics—*identity change*

Keywords
Identity, Education, Stereotype, Machine Learning

1. INTRODUCTION
Women are dramatically underrepresented in computing classrooms and careers [13]. A variety of causes for this underrepresentation have been proposed, including the "experience gap" [9, 2], and stereotypes of computing as "geeky" and boring [4]. One factor in students' sense of fit with the major is their sense of belonging [15], or their sense of identity as the kind of person who fits in.

In recent years, a number of programs have been developed to provide young women experience with computing [3]. These programs range from one-day workshops to longer summer programs. In general, these programs attempt to increase girls' experience with computing in order to decrease the experience gap and to overcome the stereotypes that computer science is dry, boring, and solitary. These programs have increased participants' interest in computing and taught skills and concepts, generally demonstrated through the use of pre- and post-survey measures [1]. However, prior studies have not investigated the attitudes of girls towards computer scientists themselves. Prior research has determined that computer scientists are stereotyped as nerdy and male [6]. However, youths may find these stereotypes to be dated, as recent media has portrayed computer science and its practitioners in a positive light. Such examples include the character Abby on the popular U.S. television show NCIS, the movie *The Social Network*, and online videos produced by Code.org. Research needs to shed light on how young women perceive computer scientists and the extent to which computing programs can change participants' perceptions.

By late childhood, students can identify traits that describe themselves [10]. By adolescence, such as our high-school aged participants, students are able to identify personal goals, motives, and values that apply to themselves. Erikson defines identity as the sense of self which is continuous and unchanged across settings [7]. As individuals mature, they become more nuanced in their understanding of themselves as social actors. They behave appropriately for different situations and roles (e.g. daughter, student, friend), and some researchers emphasize the effect of roles and situations on identity [11]. However, people generally maintain a consistent self-attribution even when performing roles differently, such as the class clown who is respectful and polite at a funeral. Nonetheless, the assumption of a social role can change people's self attributions.

Therefore, it is plausible that students who participate in an immersive activity may affiliate more closely with the domain as the result. While measures of identity are subject to the person's perception of themselves and the situation, it is possible that we may measure underlying change. Students in an intensive computing setting may be more likely to find computer science traits salient than non-stereotypical identity traits, and recall them more readily. In an all-female setting where computer science is valued by authority figures, such as teachers, as well as by peers, the potential for social sanction for identifying with a stereotypically-divergent identity is much lower than it might be in a different setting, such as a mixed-gender school or athletic competition.

While we feel this to be important, measuring student perception is difficult. For example, the "Draw a Scientist" test [5] allows for open-ended expression, but interpreting the results is both time consuming and subjective. In addition, it could be that students draw a stereotypical scientist even if they do not believe that stereotype. Further, it does not allow for comparison between students' perception of themselves and their perception of a scientist. We intend to improve this measure in a way that better reveals the intended construct.

GenderIT '15 April 24 2015, Philadelphia, PA, USA

ACM 978-1-4503-3596-6/15/04 ...$15.00.
http://dx.doi.org/10.1145/2807565.2807711

2. METHOD

2.1 Participants

This study took place as part of a summer program for high school girls, ages 15-17. Participants ($N = 162$) applied to take part in the program and were chosen based on an essay, teacher recommendation, and grades. Although girls had to be interested enough to attend, there was no expectation of prior experience with computer science.

2.2 The Program

The program took place in eight locations across the United States, with each location enrolling approximately 25 students. Students attended the program 7 to 8 hours a day, five days a week, for eight weeks. They were taught by a computer science teacher and assisted by one or two course assistants. Various guest speakers visited over the course of the program. The curriculum was designed to teach the girls a variety of computer science topics, including programming, robotics, and web design, all an introductory level. The program culminated with an open-ended project where small groups of students used what they had learned to design a technological solution to a problem they had identified.

2.3 Design and Procedure

Data were collected twice, at the beginning and end of the program. As an introductory activity, students completed a survey about their interests and experiences with computing. At the beginning of the survey, they were presented a page with 30 empty boxes with the title, "Describe Yourself" and the description, "Spend approximately 1 minute and list all the adjectives or phrases you can think of to describe yourself, such as "athletic," "creative," or "likes math." Please put each word or phrase in its own box." They then responded to the rest of the survey which had questions about their plans for the future, computing, and family support; this took approximately 45 minutes. At the conclusion of the survey, they were prompted to describe a computer scientist with the description, "Spend approximately 1 minute and list all the adjectives or phrases you can think of to describe a computer scientist, such as "athletic," "creative," or "likes math." Please put each word or phrase in its own box." A current version of this tool can be viewed at `http://awesome.stanford.edu/words`.

Data were collected again during the final days of the program. The prompts were identical to the initial survey; the parts of the survey in between included questions about students' experience in the program rather than prior computing experience, but was of similar duration.

Upon receiving the data, we performed a spell check using MS Office. In almost all cases, the intended word was obvious, but if we had any doubt, we did not alter the word (e.g. "Jonatic," which is a Jonas Brothers fan, remained unchanged.) Ten students who completed the pre-survey did not complete the post-survey, and were excluded from any comparison analysis.

2.4 Analysis

2.4.1 Perception of Computer Science

We examined two dimensions of participants' perception of computer scientists. First, we investigated how positively participants' view computer scientists, a measure we refer to as "sentiment". Second, we investigated how closely participants' perception of computer scientists matches widely-held but oversimplified images or ideas, a measure we refer to as "stereotype". We leveraged robust machine learning algorithms to measure these traits.

Stereotype	Anti-Stereotype
Smart	Passionate
Intelligent	Fun
Determined	Funny
Likes Science	Cool
Hard Working	Curious

Table 1: Most common stereotypes and anti-stereotypes (excluding words in the prompt)

Sentiment Analysis. The Natural Language Processing community has produced a substantial amount of research on Sentiment Analysis [14]. Models are trained on large datasets extracted from across the Internet to determine if a word has positive or negative connotations. Sentiment is scored on a scale from -1, very negative to +1, very positive with 0 meaning neutral. For example "intelligent" has a positive sentiment (0.9) and "sickly" is negative (-0.5). Contemporary models are able to achieve high accuracy on predicting word and short phrase sentiment. The model that we use was trained by AlchemyApi using a dataset of 200 billion words and is especially adept at "noisy" data (e.g. words with slang, misspellings and idioms) [1]. We define the sentiment of a set of student words to be the average sentiment of each of the users' phrases:

$$S(W) = \frac{\sum_{p \in W} \delta(p)}{|W|}$$

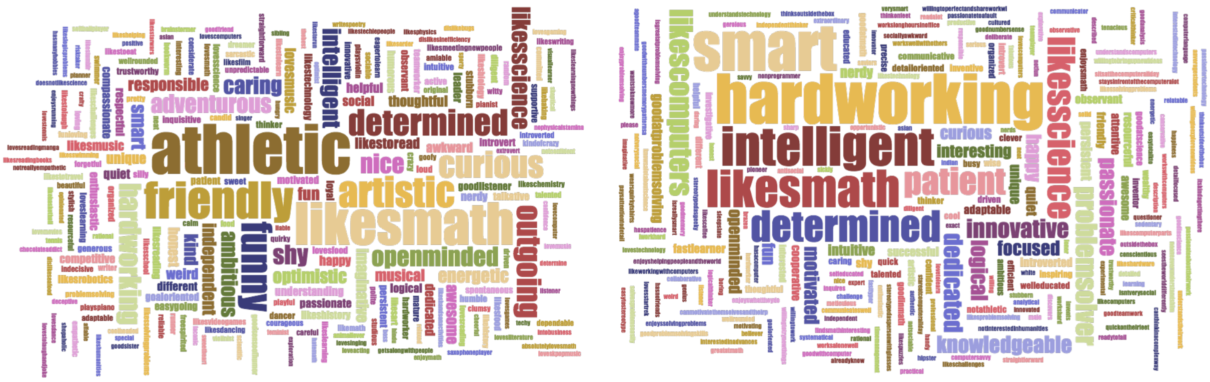
Where $S(W)$ is the sentiment of the collection of phrases W, and $\delta(p)$ is the sentiment generated by AlchemyApi for phrase p.

Stereotype Analysis. To our knowledge, a standard measurement of phrase stereotype does not exist. So we used the same intuition behind sentiment analysis to generate a measure of the degree to which a phrase conforms to the computer science stereotype. We selected the 100 most popular terms to describe a computer scientist, blind to pre/post prompt. These 100 phrases accounted for 59% of user phrases. We scored the phrases with a number +1 for stereotypical and -1 for anti-stereotypical. For example "collaborative" and "artistic" were given scores of -1 and "serious" and "likes-science" were given scores of +1. See table 1 for the most common stereotypical and anti-stereotypical terms. We then used phrase similarity measures to propagate stereotype labels to similar words [12] [2]. When we were not confident whether a phrase was stereotypical or not, it was given a neutral score of 0. Given a stereotype score for each phrase, we calculated the stereotype score in the same manner as the sentiment score.

2.4.2 Computer Science Identity

Another perspective into the attitudes of girls towards computer science is to observe the similarities and differences between the words that they use to describe themselves and computer scientists. To measure the similarity between "self" and "computer science" descriptions, we computed the Jaccard Similarity Index, which is the ratio of the number of words in common between the two sets divided by the total unique words in the two sets. A score of zero indicates that no adjectives were common between the two sets, whereas a score of 100 indicates the sets are identical.

[1]http://www.alchemyapi.com/api/
[2]https://code.google.com/p/word2vec/



(a) (b)

Figure 1: Tag cloud of the words students used to describe (a) self and (b) computer scientists on the first day of the program. Word size is proportional to popularity. The prompt word "creative" which was the most used in all descriptions, was not included.

3. RESULTS

There were 971 unique phrases students used to describe themselves and 740 unique phrases students used to describe computer scientists. Figure 1 shows the most common words that students used to describe themselves and computer scientists at the beginning of the program. In describing themselves, a paired t-test revealed no significant difference in the number of adjectives used at the beginning ($M = 8.9$, $SD = 4.47$) to the end ($M = 9.16$, $SD = 4.46$), $t(147) = -0.89, p = 0.37$. However, in describing a computer scientist, there was a significant difference in the number of adjectives used between the beginning ($M = 6.02$, $SD = 2.42$) and the end ($M = 6.91$, $SD = 2.95$), $t(147) = -3.25, p = 0.001$.

In addition, girls changed how they described computer scientists, as shown in Figure 2. Pre-survey descriptions were more stereotyped ($M = 0.41$, $SD = 0.44$) compared to post-survey ($M = 0.09$, $SD = 0.42$), which is a significant difference (two-tailed bootstrap, $p < 0.0001$). Also, descriptions were significantly more positive, from pre ($M = 0.75$, $SD = 0.37$) to post ($M = 0.89$, $SD = 0.19$), (two-tailed bootstrap, $p < 0.001$). By comparison, girls expressed more positive sentiments about themselves as well, from pre ($M = 0.76$, $SD = 0.40$) to post ($M = 0.85$, $SD = 0.283, p = 0.002$).

At the end of the program, the girls used almost twice as many common adjectives in their descriptions of selves and computer scientists than they did at the beginning of the program, as shown in Table 2. The Jaccard similarity index between self and computer scientist phrases significantly increased from 8.00 ($SD = 0.59$) to 13.32 ($SD = 9.38$), (two-tailed bootstrap, $p < 0.0001$). More girls had at least one common adjective rather than a few girls having many more common adjectives–at pre-survey, 58.8% of participants had a non-zero Jaccard index, at post, 79.3% had at least one common adjective.

We found that the overlap between student's post description of computer scientists and their pre description of self (7.0) was lower than the overlap between their post description of self and their pre description of computer science (8.5), also shown in Table 2, This is evidence that the changing perception of self drove the convergence between self descriptions and CS descriptions.

	preCs	preSelf	postCs	postSelf
preCs	-	**8.0**	19.2	8.5
preSelf	8.0	-	7.0	19.2
postCs	19.2	7.0	-	**13.3**
postSelf	8.5	19.2	13.3	-

Table 2: Mean Jaccard Similarity between sets of responses.

4. DISCUSSION

We asked high school students to describe themselves and a computer scientist both before and after an eight week computer science program. In describing themselves, they used on average, nine adjectives both before and after. In describing computer scientists, from the beginning of the program to the end, participants were more positive, less stereotypical, and on average they provided an additional adjective. We view this as evidence that they have a better understanding of what is a "computer scientist."

We can imagine how this could come to be. In the beginning, the participants may have had a vague notion of a computer scientist, and may not have had any particular person in mind when they were describing a computer scientist. Even if a girl had a parent who is a computer scientist, that parent would play the role of Mom or Dad who happens to do computers while she is at school. However, at the end of the program, they have had many interactions with people who they primarily identify as computer scientists. The instructors, the teaching assistants, and guest speakers would all interact primarily in that role. We have some evidence that students may have been thinking of a specific person at post-survey in that one of the largest increases in adjectives was the word "helpful."

With this possible mechanism in mind, we still find a convergence between phrases used to describe the self and phrases used to describe a computer scientist from the beginning to the end of the program. Independent of the technical skills they learned over course of the program, these participants saw themselves as more similar to a computer scientist. In examining professionals making transitions in the workplace, Ibarra found that one task was to observe role models to identify potential identities, and another was to experiment with a provisional self [8]. We suggest that this converging list of descriptive phrases is preliminary evidence of both.

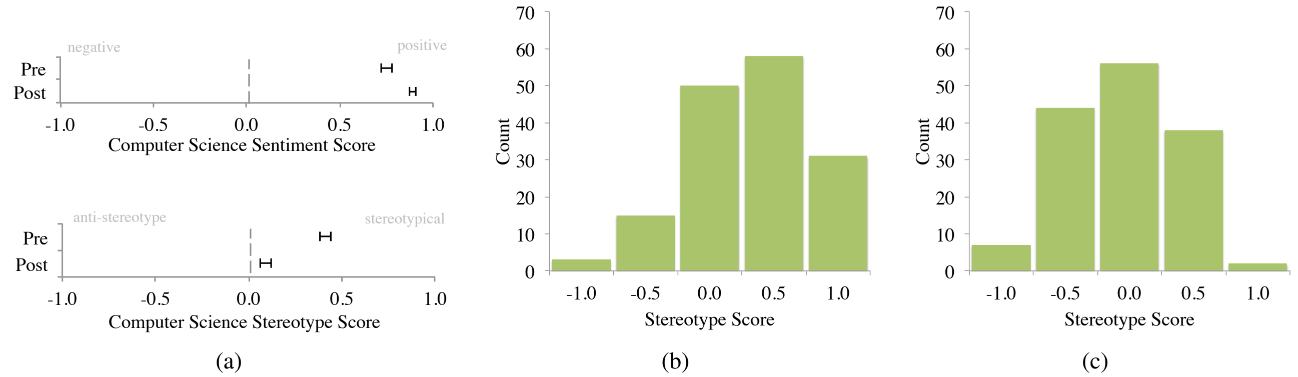


Figure 2: The change in the perception of Computer Scientists shown in (a), which is the difference in means for sentiment and stereotype scores. The histograms of stereotype scores for (b) the pre-test and (c) the post-test.

5. CONCLUSION

Our easy to administer and relatively unobtrusive measure has shown that participants of one particular program view themselves more similarly to computer scientists at completion. Students self-selected to attend this program, and had generally positive attitudes throughout. With regard to this positive sentiment, we are pleased to report that we did not find a ceiling effect with a group who would be likely to demonstrate one.

One limitation of this work is that our team labeled the most common words with our own contemporary perception of stereotype. We attempted to weight the words as stereotypical, neutral, or astereotypical independent of whether they were positive (smart) or negative (geeky.) However, we may not be hip to the jive of what the kids are stepping in these days. (And that sentence is almost certain proof that we are not always picking up what they are putting down.) Therefore, we must expand and trim the lexicon of stereotypical words as language evolves. It is not clear that the same stereotypical words will be stereotypical five years from now. This will be an area of focus for us.

Another step we intend to take is to suggest the use of this measure with other programs that are less time intensive. For example, we might consider comparing a required computer science class to a non-required one. We hypothesize that the stereotypical measures and the sentiment measures may change differently in these two courses.

Without a doubt, we need more women computer scientists. We hope to contribute by providing a measure that gives formative feedback to programs and classrooms that have that aim, making this larger endeavor more successful.

6. REFERENCES

[1] C. Ashcraft, E. Eger, and M. Friend. Girls in IT : The Facts. Technical report, National Center for Women in Information Technology, 2012.

[2] B. Barron. Learning Ecologies for Technological Fluency: Gender and Experience Differences. *Journal of Educational Computing Research*, 31(1):1–36, Mar. 2004.

[3] A. Bruckman, M. Biggers, B. Ericson, T. Mcklin, J. P. Dimond, B. Disalvo, M. Hewner, L. Ni, and S. Yardi. âĂIJ Georgia Computes!âĂİ: Improving the Computing Education Pipeline. In *SIGCSE '09*, pages 86–90, 2009.

[4] L. Carter. Why students with an apparent aptitude for computer science don't choose to major in computer science. *Proceedings of the 37th SIGCSE technical symposium on Computer science education - SIGCSE '06*, page 27, 2006.

[5] D. W. Chambers. Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, 67(2):255–265, 1983.

[6] S. Cheryan and V. C. Plaut. Explaining Underrepresentation: A Theory of Precluded Interest. *Sex roles*, 63(7-8):475–488, Oct. 2010.

[7] E. Erikson. *Childhood and Society*. Norton, 1963.

[8] H. Ibarra. Provisional selves: Experimenting with image and identity in professional adaptation. *Administrative Science Quarterly*, 44:764–791, 1999.

[9] J. Margolis and A. Fisher. *Unlocking the Clubhouse: Women in Computing*. The MIT Press, 2003.

[10] D. P. McAdams. The psychological self as actor, agent, and author. *Perspectives on Psychological Science*, 8:272–295, 2013.

[11] A. R. McConnell. The multiple self-aspects framework: self-concept representation and its implications. *Personality and social psychology review*, 15(1):3–27, Feb. 2011.

[12] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119, 2013.

[13] NCWIT. By the Numbers. Technical report, National Center for Women and Information Technology, 2014.

[14] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135, 2008.

[15] N. Veilleux, R. Bates, D. Jones, J. Crawford, C. Allendoerfer, and T. Floyd Smith. The relationship between belonging and ability in computer science. *Proceeding of the 44th ACM technical symposium on Computer science education - SIGCSE '13*, 6:65, 2013.