Choosing Computer Science:
Women at the Start of the Undergraduate Pipeline

Ken Yasuhara

yasuhara@cs.washington.edu
Department of Computer Science and Engineering
University of Washington
Seattle, WA

1 Introduction

The "shrinking pipeline" is a common metaphor for the underrepresentation of women in computer science (CS), an increasingly well-known (if not well-understood) phenomenon. The further one progresses in CS academia—from undergraduate study to graduate study to faculty rank—the fewer women there are. (For a comprehensive discussion of the underrepresentation of women in CS, see Gürer and Camp.[13]) At the undergraduate level in the U.S., CS is the only science, technology, engineering, and mathematics (STEM) field whose gender gap has *widened* during the last two decades.[11] In the U.S., only 28% of Bachelor's degrees in computer and information sciences went to women in 2002, down from a high of nearly 40% in the mid 1980s (Figure 1). As in past years, research departments are faring worse; in U.S. and Canadian Ph.D.-granting departments, 18% of Bachelor's degrees in computer science and engineering went to women in 2003.[25]

This paper reports early results of a survey- and interview-based study focusing on the beginning of the undergraduate pipeline in CS. Employing a grounded theory-style method, we investigate gender differences in how pre-major undergraduates are attracted to (or repelled by) the CS major. By also investigating how students conceive of CS as a discipline, culture, and career area, we not only ask the students the question, *Why or why not computer science?*, but also discover how they *understand* the question.

We begin by discussing relevant underrepresentation research in Section 2. Section 3 presents research questions and our methods. Survey analysis results in Section 4 are interpreted and discussed in Section 5, and the paper concludes with a discussion of future work.

Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition

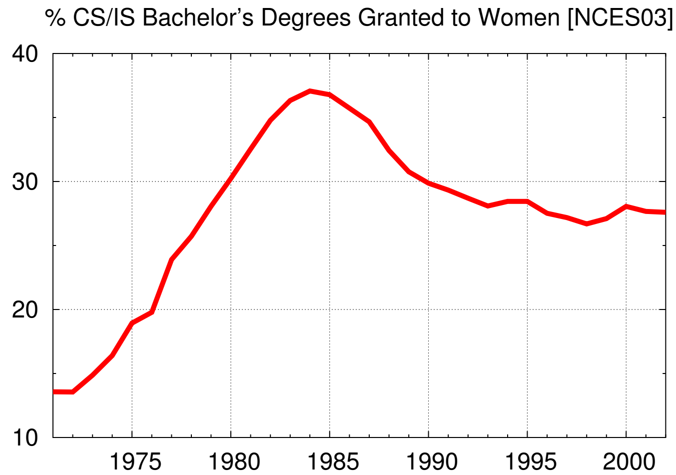

Figure 1: The gender gap in undergraduate CS has widened since the 1980s.

2 Related work

Unfortunately, there are few recent studies on the reasons for the underrepresentation of women in undergraduate computer science. Whatever the reasons, it is clearly *not* because women are any less capable intellectually. Studies of undergraduates in CS and other STEM fields show that many of the women who leave these majors are academically as successful as the men who persist.[15,22] This suggests that we should examine non-cognitive factors influencing women's persistence in CS, including affective factors and pre-college experiences. For instance, Margolis & Fisher's Carnegie Mellon study[15] highlighted self-confidence and pre-college experience with computer programming as significant and related factors. Cohoon's Virginia-wide study showed how persistence and self-confidence are related to other factors, such as mentoring and peer support.[7]

Most studies focus on the reasons women leave CS programs. These studies sample from the population of students who have committed (at least initially) to the CS major. This population is relatively easy to access, and their perspectives are very valuable for understanding why students switch out of the CS major. To the extent that most major-switching occurs during the second year,[15,20] these studies illuminate the experiences of CS majors who drop out around the middle of the undergraduate pipeline.

We stand to learn as much, however, from a distinct population: those students who never choose the CS major in the first place, early in their undergraduate careers. There are usually more practical obstacles (*e.g.*, access, sampling) to studying this population, compared with studying students who have declared CS. However, their perspectives are valuable for understanding the very start of the undergraduate pipeline. Seymour & Hewitt's landmark study of STEM majors found that persisters and nonpersisters differed more in their reasons for *choosing* a STEM major than in the concerns that eventually caused nonpersisters to leave.[20] This finding underscores the importance of research at the start of the undergraduate pipeline.

3 Studying the pre-major perspective

This study aims to extend the literature on the undergraduate pipeline by examining the following research questions:

- How do perceptions of CS affect pre-major undergraduates' interest in the CS major?
- How do these perceptions and effects vary with gender?

3.1 Qualitative, inductive method

Based on past research, we might expect certain factors (self-confidence, prior programming experience) to affect interest in the major. However, rather than presuppose the relevance of a specific set of factors, we opted for a grounded theory-style method,[6] in which we ask students open-ended questions about perceptions of CS and interest in the major, expecting themes or patterns to emerge in the data.

Although this inductive, "bottom-up" method is laborious, advantages of comprehensiveness and richness make it appropriate for our research questions. Literature in this area is limited, and the relevant factors might vary with local context.[7,15,20] Our method favors discovering or validating a wide range of factors affecting interest in CS. By asking students to discuss these factors in their own words, rather than in terms of hypotheses framed by the researcher, we can hope for a richer, possibly more authentic understanding of their perspectives, including relationships among the factors.

3.2 Survey and analysis

In the first of two stages in this study, we administered anonymous surveys to pre-major undergraduates. The main goals of the survey were to discover the range of factors affecting choice of the CS major and which of those factors are more common among women and men. A set of three broad, open-ended questions about the CS major comprised the survey's core and are discussed more in Sections 4.2 and 4.3. We analyzed the survey responses and have identified key themes to examine in more detail in the second stage of the study, which will consist of semi-structured interviews. This paper reports findings from the first, survey-based stage.

Although most of the research effort went into transcription and analysis of the survey responses, the survey questions themselves were carefully worded and ordered (*e.g.*, to minimize stereotype threat[21]), and the instrument was reviewed by the university's human subjects division. In addition to asking directly about students' interest in the CS major, we elicit their personal conceptions of the field of computer science. This allows us to avoid the assumption that a student's definition of "computer science" resembles a computer scientist's definition. These student conceptions, regardless of accuracy, inform students' choices of major. To accurately interpret students' responses on what they find (un-)interesting about CS, we put them in the context of these conceptions.

Survey responses were transcribed and coded, mapping each response to one or more themes. Each open-ended question had a set of themes that was induced from the responses in a first pass of coding. Themes were added and refined through a process of comparison with each new response. Once the themes stabilized, a second pass of coding ensured each response was compared against the complete set of themes. For verification and fine-tuning, one third of the responses (randomly selected) and descriptions of the themes were given to a second researcher for coding.

3.3 Institutional and curricular context

The survey was completed in 2004 spring by students enrolled in the first of two introductory computer programming courses (referred to here as CS 1 and 2, respectively) at a large, public university classified as "Doctoral/Research Universities—Extensive" by the Carnegie Foundation.[17] With instructor permission, the author administered the survey in class during the term's first week, in an effort to capture the students' perspectives prior to exposure to the introductory course.

The department offers two undergraduate majors: Computer Science and Computer Engineering. Prerequisites are very similar for both majors and include calculus, science, and the CS 1 and 2 courses. The required core courses in both majors also largely overlap, with Computer Science majors having slightly more freedom to choose topics courses later in their studies. Therefore, this paper discusses the two majors together and refers to them as "CSE."

CS 1 is titled "Computer Programming I," has no formal prerequisites, and is listed with no requirement of prior experience with computer programming. (The department does explicitly acknowledge, however, that such experience can be helpful.) The course covers programming basics using Java, with topics including control flow, iteration, arrays, searching, sorting, and object orientation. CS 1 is required for students who wish to major in computer science. However, being both a pre-major and a service course, it is required for all pre-engineers, and only a small fraction of CS 1 enrollees go on to become CSE majors. Therefore, while we do not assume CS 1's enrollment is a representative sample of the general university population, we can expect to get perspectives other than those of students who have committed to studying CSE. Furthermore, from the standpoint of recruiting CSE majors, the students in CS 1 are "low-hanging fruit," in the sense that the department can use CS 1 to pique their interest in majoring.

In order to keep class sizes small for majors, the department restricts CSE enrollment through competitive admissions. Student must apply to become CSE majors after completing prerequisite coursework, usually in the second or third year. The same admissions committee makes decisions on applications for both the CS and CE applicants according to a single set of criteria, including academics, personal attributes, and an essay. Roughly one third to half of the applicants have been accepted in recent years, with women and men accepted at comparable rates.

At this university, the "shrinking pipeline" effect is visible even at the level of the CS 1–2 introductory sequence; female representation drops from CS 1 enrollment (25–30%), to CS 2 enrollment (about 20%), and to the CS applicant pool (15–20%). This study is motivated in part by the large number of prospective female majors lost during or after the first course.

4 Survey results

4.1 Sample profile

About three hundred students completed the two-page survey. Of the three hundred respondents, we focused analysis on the 205 who were first- or second-year students. These students could reasonably consider applying to major in CSE. Fewer than one fifth of the students indicated interest in CSE. Women were twice as likely to indicate interest in CSE than men (Figure 2). Because of the small number of women in the sample (51, 25%), this meant 40% of students interested in CSE were women.

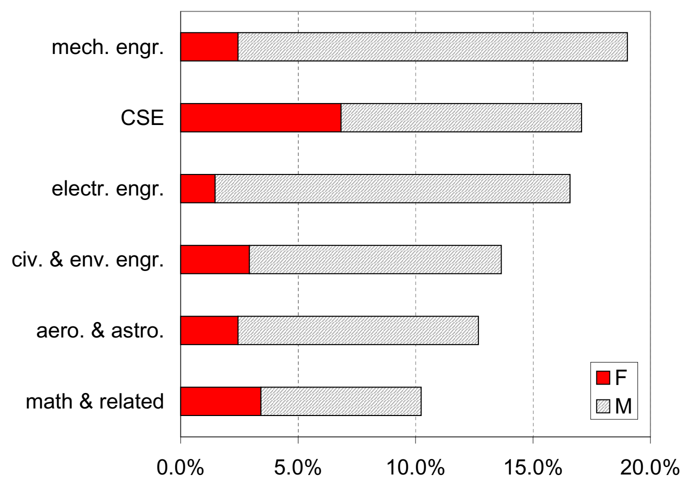

Figure 2: Majors and selected categories of majors by level of student interest: Bar length shows fraction of full sample who indicated interest in each major, with bar division showing gender ratio among students interested in the major. Figure only shows majors listed by at least 10% of sample. Students were allowed to list as many majors as they wished.

Two thirds of the students reported having no prior programming experience, with no significant gender difference (Figure 3). Among the remaining third, there were gender differences for certain kinds of experience. There was no significant gender difference with respect to programming experience from high school and undergraduate courses, but women were less than half as likely to have had self-taught or other extracurricular experience. (Due to small and/or uneven marginal frequencies, statistical significance was determined by the Fisher exact test, $p \leq 0.05$.)

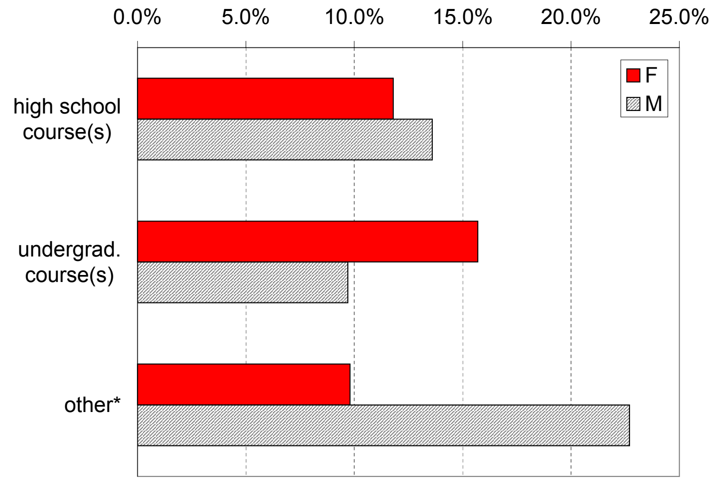

Figure 3: Programming experience by type and gender: Experience types are high school course(s), undergraduate course(s), and extracurricular or self-taught (other) experience. (*statistically significant gender difference, $p \leq 0.05$)

4.2 Computer science according to the pre-major

One of the open-ended survey questions asked students to try to define "computer science" in their own words. Although the question was asked later on the survey, we discuss the responses first, to provide context for the following discussion of factors related to interest in CS. Of the 205 students, 150 wrote answers to this question. Most of the responses were brief (average of fourteen words), and for the most part, women and men answered similarly.

The most common theme was programming, mentioned in almost half of the responses. Slightly less than one third of the responses mentioned applications of computing (*e.g.*, to solve problems, to automate tasks). About one quarter of the responses said CS is about how computers work. 80% of responses matched at least one of these three themes. About one tenth of responses mentioned the following themes: technology and/or its advancement, hardware or the physical composition of computers, and design. Themes that matched less than 5% of responses included mathematics, creativity, engineering, information or data, fixing computers or software.

Only two themes appeared with significantly different frequency in women's and men's responses. The first was creativity, which matched none of the women's responses but 10% of men's. The opposite was true of characterization of CS as a broad field (*i.e.*, encompassing many subfields and/or having applications in many areas), which matched 9% of women's responses but only 1% of men's.

4.3 Reasons for more/less interest in CS

The survey also included a pair of questions on aspects of the CS major that make it more/less interesting. We asked a balanced pair of questions rather than assume a polarized opinion, which might also have countered any implication that the students were expected to be interested in the major (or vice versa). Of the 205 students, 150 wrote answers to these questions, and again, responses were brief, averaging eleven words.

For both women and men, the most commonly cited interesting aspects of CS were programming and applications of computing, matching about one fifth of responses. Next came enjoyment of computers/technology and rapid growth/advancement of computing, each matching about one tenth of responses. Just under one tenth of responses mentioned career advantages afforded by studying CS, *e.g.*, job availability and salary. 60% of the responses (excluding 23 respondents who simply indicated disinterest in CS rather than answer the question) matched at least one of these top five themes.

Certain themes ranked higher among women's responses. Creativity and logic/math each matched 12%, and intellectual challenge and the practical value of computing skills/knowledge each matched 9%. Of these themes, the gender differences were statistically significant only for creativity and logic/math.

As for aspects that make CS less interesting, women and men were more in agreement, collectively, with none of the themes exhibiting significant gender differences. One fifth of responses cited excessive workload, and one sixth mentioned programming. At least one tenth of responses matched each of the following four themes: asocial, sedentary lifestyle; (intellectual) difficulty of the major; dullness/tediousness; and competitiveness (*e.g.*, major admissions, course atmosphere). 65% of the responses (again excluding non-answers) matched at least one of these top six themes.

5 Discussion

The women in our sample were just as likely as the men to have had programming experience from high school courses. However, prior studies of programming experience[15,19,24] and Advanced Placement CS exam statistics[8] indicate a gender gap in the high school pipeline. The likely reason our sample did not exhibit this gender gap is the self-selection of the women enrolled in CS 1. The higher likelihood of women to be interested in the CS major is additional evidence for self-selection. If indeed women with little or no programming experience are opting not to enroll in CS 1, it might be because of the course's reputation for not being novice-friendly, in spite of the department's official advertisement that the course has no prior experience requirement. A number of students mentioned CS 1's reputation as an exteremely challenging course, and two responses specifically mentioned the challenge posed to novice programmers.

More in keeping with expectations, the gender gap in extracurricular and other programming experience might reflect the lower rate of home computer ownership/access[3,9,14,15,19,22] or less comfort or confidence with computing among women.[15,24] Avoiding extracurricular programming might also be a reaction to the associated "geek" stereotype and gender socialization.[15,16]

The prominence of programming in students' definitions of CS is consistent with Blum & Frieze's account of Carnegie Mellon University CS majors who initially equated CS and programming.[4] Computer scientists might be dismayed at how narrowly most students construe their field, with only a handful of students in this sample associating CS with mathematics, creativity, engineering, and information, and fewer still discussing (interpersonal) communication.

The tendency for students to miss the significance of mathematics in CS might represent an untapped opportunity to recruit more women. In the U.S., mathematics has nearly achieved gender parity at the undergraduate level. CS might similarly become more accessible to a wider audience, if introductory courses can more effectively convey connections between CS and mathematics. This strategy seems promising considering women in our study were especially attracted to the mathematics and logic in CS, a finding corroborated in a recent national study of CS majors.[23]

In addition to mathematics, our findings suggest that an introductory course with room for creativity might recruit more women. The attraction of women to creative aspects and applications of CS might explain the success of Georgia Tech's multimedia-focused, gender-friendly alternative to conventional introductory CS.[18]

Based on prior research, we expected applications of computing and the role of computing in society to rank higher as attractors to CS for women,[12,15,23] but these aspects were cited equally frequently by men. As with high school programming experience, this might be an effect of self-selection. Notably, our survey data provides little indication of how important these factors are to women and men; relative weighting of factors remains to be studied in the interviews. Another possible explanation is the ambiguity of the word "application," which can refer to software or a way in which computing can be applied in a domain or problem. In some responses, context was unhelpful in determining the intended meaning. This natural limitation of surveys can be overcome by interviews, and we hope they will elicit more detailed student perceptions about the applications of computing and their potential significance to women.

Student concerns about workload should prompt CS instructors to examine whether their introductory courses present a reasonable level of challenge for novice programmers. Consistent with our findings, many CS students perceive their introductory courses to be inappropriately geared for students who already have some programming experience.[5,15] At least one CS 1 study reported that most of the top grades were achieved by students with moderate precourse familiarity with programming concepts.[19]

6 Future work

In addition to another round of surveying, we expect that interviews will provide for a more refined and confident analysis of gender differences in attraction to the CS major. A larger survey sample might make disaggregation of more specific subsamples feasible, *e.g.*, allowing us to compare pre-major women who are interested in CS with those who are not. With the ability to follow-up and shift focus "on the fly," semi-structured interviews will provide the details and context missing from the brief, written responses in the survey. Approaches we are considering to better understand how students conceive of CS include asking them to imagine and describe a successful CS student or a professional computer scientist's typical day of work.

We hope interviews will also help disambiguate themes that appeared frequently in the surveys. Intriguingly, many students said programming made CS more interesting, while many others said the opposite. This, coupled with the frequency with which students mentioned programming in their definitions of CS, suggests programming is a complex construct. Perhaps students are thinking of different parts of the programming process or different aspects of the activity of programming. To the extent that programming can be a means to an end, we might expect the task accomplished through programming to affect how students perceive and value the activity. By eliciting details and context during interviews, we hope to gain a more complete, nuanced understanding of the role of programming in students' perceptions of CS. Other themes warranting closer examination include creativity, workload, and logic/math.

Looking beyond this study to implications on teaching practice, we might attempt validating our findings by using them to design an alternative or supplemental introductory CS course. Informed by common preconceptions of CS, such a course might, for example, be more effective at conveying a more complete and accurate view of the field than conventional CS 1 courses. With women representing 40% of CS 1 students considering majoring in CSE (a far larger fraction than in the major), this course represents both a severe pipeline leakage and a critical opportunity to encourage more women to study CS. We expect that keeping to these practical, long-term goals will continue to help focus and scope the study.

Finally, without replicating this or similar studies at other institutions, we cannot assume our findings generalize across institutions. As mentioned in Section 3, with wide variation in local culture, context, and population, we should not assume that the dynamics of the gender gap in one department are necessarily similar to those of another. Resources and cooperation permitting, we are considering at least a survey-based study of analogous populations at other institutions. With mounting evidence that the CS gender gap varies widely internationally,[1,2,10] we are particularly interested in including CS departments outside the U.S.

7 Conclusions

This study extends the CS-specific literature on the underrepresentation of women by employing a grounded theory-style method and focusing on students who are unlikely to major in CS but are good candidates for recruitment. At least at the studied institution, the shrinking pipeline effect is evident even in the first terms of coursework. Many of our early findings validate prior research highlighting stereotypes and workload as negative factors. Our findings suggest that emphasizing the breadth of CS, particularly aspects of mathematics and creativity, can counter the disproportional loss of women early in the undergraduate pipeline. As the second, interview-based stage of our study proceeds, we hope to gain a more detailed, contextualized understanding of how women perceive and choose to study CS.

Acknowledgments. The author thanks the students and instructor for their cooperation and J.M. Cohoon, R. Adams, V. Sazawal, and E. Jaślikowska for their advice and assistance. This work was supported in part by funding from the Center for the Advancement for Engineering Education.

References

[1] Joel C. Adams, Vimala Bauer, and Shakuntala Baichoo. An expanding pipeline. In Scott Grissom, Deborah Knox, Dan Joyce, and Wanda Dann, editors, *Proceedings of SIGCSE'03: the* 34^{th} *SIGCSE technical symposium on Computer Science Education*, pages 59–63, Reno, NV, USA, February 2003. ACM Press.

[2] Susan Bergin and Ronan Reilly. Programming: Factors that influence success. In *Proceedings of SIGCSE'05: the* 36^{th} *SIGCSE technical symposium on Computer Science Education*, 2005.

[3] Danielle R. Bernstein. Is teaching computer science different from teaching other sciences? In *Proceedings of the Thirteenth Annual Eastern Small College Computing Conference*, October 1997.

[4] Lenore Blum and Carol Frieze. As the culture of computing evolves, similarity can be the difference. `http://www-2.cs.cmu.edu/~lblum/PAPERS/TheEvolvingCSCulture.pdf`, 2004.

[5] Eileen D. Bunderson and Mary Elizabeth Christensen. An analysis of retention problems for female students in university computer science programs. *Journal of Research on Computing in Education*, 28(1):1–18, 1995.

[6] Amanda Coffey and Paul Atkinson. *Making sense of qualitative data*, chapter 2, pages 26–53. Sage Publications, 1996.

[7] J. McGrath Cohoon. Toward improving female retention in the computer science major. *Communications of the ACM*, 44(5):108–114, May 2001.

[8] Advanced Placement Program National Summary Report. `http://www.collegeboard.com/ap/library/state_nat_rpts_02.html`, 2002.

[9] A. Craig. Peer mentoring female computing students. In *Proceedings of ACSE 1998*, 1998.

[10] Larisa Eidelman and Orit Hazzan. Factors influencing the shrinking pipeline in high schools: A sector-based analysis of the israeli high school system. In *Proceedings of SIGCSE'05: the* 36^{th} *SIGCSE technical symposium on Computer Science Education*.

[11] National Center for Education Statistics. Digest of education statistics. `http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005025`, 2003.

[12] Irene F. Goodman, Christine M. Cunningham, Cathy Lachapelle, Meredith Thompson, Katherine Bittinger, Robert T. Brennan, and Mario Delci. Final report of the women's experiences in college engineering (WECE) project, April 2002.

[13] Denise Gürer and Tracy Camp. Investigating the incredible shrinking pipeline for women in computer science. `http://women.acm.org/documents/finalreport.pdf`, June 2002.

[14] Andrea Jepson and Teri Perl. Priming the pipeline. *SIGCSE Bulletin*, 34(2):36–39, June 2002.

[15] Jane Margolis and Allan Fisher. *Unlocking the Clubhouse*. MIT Press, 2002.

[16] C. Dianne Martin. Draw a computer scientist. *SIGCSE Bulletin*, 36(4):11–12, 2004.

[17] Alexander C. McCormick, editor. *The Carnegie Classification of Institutions of Higher Education, 2000 Edition*. Carnegie Foundation for the Advancement of Teaching, 2001.

[18] Lauren Rich, Heather Perry, and Mark Guzdial. A CS1 course designed to address interests of women. In Dan Joyce, Deborah Knox, Wanda Dann, and Tom Naps, editors, *Proceedings of SIGCSE'04: the* 35^{th} *SIGCSE technical symposium on Computer Science Education*, pages 190–194, Norfolk, VA, USA, March 2004. ACM Press.

[19] Marian Gunsher Sackrowitz and Ann Parker Parelius. An unlevel playing field. In John Impagliazzo, Elizabeth Adams, and Karl J. Klee, editors, *Proceedings of SIGCSE'96: the* 27^{th} *SIGCSE technical symposium on Computer Science Education*, pages 37–41, Philadelphia, PA, USA, February 1996. ACM Press.

[20] Elaine Seymour and Nancy M. Hewitt. *Talking About Leaving*. Westview Press, 1997.

[21] Claude M. Steele. Thin ice: "stereotype threat" and black college students. *Atlantic Monthly*, 284(2):44–54, August 1999. Available at `http://www.theatlantic.com/issues/99aug/9908stereotype.htm`.

[22] A. Christopher Strenta, Rogers Elliott, Russell Adair, Michael Matier, and Jannah Scott. Choosing and leaving science in highly selective institutions. *Research in Higher Education*, 35(5):513–547, 1994.

[23] Heather K. Tillberg and J. McGrath Cohoon. Attracting women to the cs major (preprint). 2004. Available at `http://curry.edschool.virginia.edu/ITattrit/Papers/RGC.pdf`.

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

[25] Stuart Zweben and William Aspray. 2002–2003 Taulbee Survey. *Computing Research News*, 16(3):5–19, May 2004.