# Exploring Factors that Influence Computer Science Introductory Course Students to Persist in the Major

Lecia J. Barker
School of Information
University of Texas
Austin, Texas 78712 USA
+1 512 471 3821
lecia@ischool.utexas.edu

Charlie McDowell
Computer Science Department
University of California
Santa Cruz, CA 95064 USA
+1 831 459 4772
charlie@soe.ucsc.edu

Kimberly Kalahar
National Center for Women & IT
University of Colorado
Boulder, CO 80309 USA
+1 303 735 6004
kalahark@colorado.edu

## ABSTRACT
This paper describes an exploratory study to identify which environmental and student factors best predict intention to persist in the computer science major. The findings can be used to make decisions about initiatives for increasing retention. Eight indices of student characteristics and perceptions were developed using the research-based Student Experience of the Major Survey: student-student interaction; student-faculty interaction; collaborative learning opportunities; pace/workload/prior experience with programming; teaching assistants; classroom climate/pedagogy; meaningful assignments; and racism/sexism. A linear regression revealed that student-student interaction was the most powerful predictor of students' intention to persist in the major beyond the introductory course. Other factors predicting intention to persist were pace/workload/prior experience and male gender. The findings suggest that computer science departments interested in increasing retention of students set structured expectations for student-student interaction in ways that integrate peer involvement as a mainstream activity rather than making it optional or extracurricular. They also suggest departments find ways to manage programming experience gaps in CS1.

## Categories and Subject Descriptors
K.3.2 [**Computer and Information Science Education**] Computer Science Education

## General Terms
Measurement, Human Factors.

## Keywords
Retention; attrition; persistence; student-student interaction; peer interaction; gender; pace; experience gap; "Student Experience of the Major"; regression analysis.

## 1. INTRODUCTION
Which factors best predict student intention to persist in an undergraduate computer science major? This study answers that question for one computer science undergraduate major using a research-informed survey and linear regression. The publicly available assessment instrument, the Student Experience of the Major Survey [4], was developed in collaboration between researchers at the Bren School of Information and Computer

Sciences at the University of California-Irvine and the National Center for Women & Information Technology. Funded by the U.S. National Science Foundation, the survey has been customized for use by several other CS departments for identifying the strongest predictors of student persistence in the major. This study presents the survey results from one institution and uses linear regression to identify the strongest predictors of intention to major in CS at that institution. The use of regression allows decision makers to select solutions with greatest potential for impact in their own department.

## 2. LITERATURE ON STUDENT PERSISTENCE
Theories of and research on student outcomes in higher education suggest that students' individual characteristics combine with and are influenced by the educational environment to produce student achievement. Astin's Input-Environment-Output model [1] has influenced most conceptual frameworks for accounting for and planning programs for student outcomes in institutions of higher education (e.g., [2][3][4][10][12][15][17] ). Factors related to student persistence generally include student background characteristics (e.g., gender, race/ethnicity, pre-college educational experiences); institutional characteristics (e.g., size, selectivity); student-faculty and student peer interaction; student satisfaction with the learning environment; and students' ability to be involved in the academic experience (both educational and social) [2][12]. At the department or program level, those factors that faculty and administrators can control are likely to be of the greatest interest; these include student-faculty and student peer interaction, student interaction with the learning environment, and student engagement with the academic community.

Studies conducted in CS are consistent with research in higher education institutional studies, though CS has features that also make it unique (e.g., presence or lack of programming experience; tendency to be unattractive to women). Factors related to persistence that have been studied (cf. [4][6][7][10]) include the positive impact of prior experience with programming (shown to be positively associated with success in introductory courses) as well as pace and workload; a negative relationship between perceived low grades and persistence (with women leaving at higher rates than men, even with the same grades); a perception of low social relevance or meaningfulness of curricula and assignments; low levels of student-faculty interaction, including faculty attitudes, feedback, encouragement, mentoring, and career advice; problems associated with student-student interaction, such as not feeling like one belongs, a heavy focus on individualized learning, or lack of access to peer support networks; issues of pedagogy, such as the positive influence of collaborative learning

on retention (e.g., pair programming); and the quality and training of teaching assistants. For students who are members of groups underrepresented in computing, a lack of role models has been hypothesized as problematic [7]. For women, loss of confidence, often due to not understanding where they stand academically with relation to peers or comparing to their excellent high school grades, can lead to attrition [7].

## 3. STUDENT EXPERIENCE OF THE MAJOR SURVEY

The Student Experience of the Major Survey was developed on the basis of the factors shown to influence persistence, both in higher education research in general and in CS research in particular. The survey was piloted and refined on two occasions at the University of California-Irvine Bren School of Information and Computer Sciences. The survey asks students about several categories of personal characteristics and their experiences in and perceptions of the learning environment (usually for a particular class early in the major). The responses to these items are then compiled into indices that can be used in regression analysis. The indices are shown in Table 1.

For the study presented here, the survey was customized for use at a large, public, research university on the west coast of the United States. The survey asked students to respond to the survey items based on experience in an introductory programming course. In addition to survey items related to the topics in Table 1, students are asked to respond to demographic questions (race/ethnicity, gender, age, year in school); major and reasons for declaring the major or not choosing CS; in which of three quarters they had taken the course; and two measures of intention to persist (feelings of belonging and likelihood of completing major).

## 4. METHOD

### 4.1 Procedure and Sample

SurveyMonkey online survey software was used for data collection and eliciting the sample. The survey link was emailed to all 294 students who had taken the introductory programming course during the 2006-7 academic year. A $5 gift certificate was offered to those who completed the survey to reduce selection bias and serve as incentive. In addition to the original email invitation, three email reminders were sent to students. The result was a 52% response rate, with 146 usable surveys. Not surprisingly, analysis showed that juniors and seniors were very unlikely to intend to major in the course (often, it was required or elective for another major). Thus, the analysis for predicting influences on student intention to persist in the major was limited to those 113 students who were freshmen and sophomores.

The largest racial/ethnic group in the sample was White, as shown in Table 2, followed by Asian and Hispanic. The "other" category represents students who chose not to report as well as students who claimed Middle Eastern or Indian Subcontinent descent. About 10% of students in the sample who reported gender were female. The variable Underrepresentation in Computing was created by combining females with male members of underrepresented groups. Students ranged from 18-21 years of age and all but one were U.S. citizens.

**Table 1: Survey Categories/Indices**

| Index | Survey Topics |
|---|---|
| Student-Faculty Interaction | Asking professor/instructor about homework questions; positive feedback, advice on career, advice for succeeding in the major, encouragement to persist in major from professor/instructor; comfort talking to professor one on one. |
| Pace/Workload/Experience | Hours per week spent on homework and lab; perception of pace; comparison of experience required for class assignments to other courses; pre-college programming experience. |
| Student-student Interaction | Asking other students about homework; feelings about opportunities to talk to other students in class; participation in study group; spending non-school time with other students from class; membership in student organizations. |
| Classroom Climate/Pedagogy | Timely feedback; rating of instruction; professor use of student names; professor treatment of showboating students. |
| Teaching Assistants | Desire for more help in lab (scheduled, unscheduled); asking TA questions about homework; frequency and comfort of asking TAs for help; rating of TA quality; TA encouragement. |
| Collaborative Learning | Beliefs that collaboration is cheating; opportunities for collaboration in class; encouragement to work together by professors. |
| Meaningful/Relevant Assignments | Interest in assignments; perceived relevance of assignments to career goals and society. |
| Racism/Sexism | Beliefs that some students are treated worse/better on basis of race/sex; frequency of observation of racial or sexist jokes. |

**Table 2: Sample Demographics**

| Race/Ethnicity | % | Gender | % |
|---|---|---|---|
| White/Caucasion | 49% | Male | 90% |
| Asian | 19% | Female | 10% |
| Hispanic | 14% | | |
| Other | 9% | **Year** | **%** |
| African American | 5% | Freshman | 72% |
| Native American/Pacific Islander | 4% | Sophomore | 28% |

| Underrepresentation in Computing | % |
|---|---|
| Not Underrepresented | 70% |
| Underrepresented | 30% |

### 4.2 Data Procedures and Analytical Method

The eight indices described in Table 1 were tested for reliability using estimations of internal consistency among the item variables comprising each index. The alpha reliability coefficient for each scale measured at .60 (Cronbach's alpha) or higher. This measure is considered acceptable for exploratory research; it should also be pointed out that the scales had only 3-6 items each, which typically results in lower alpha coefficients. For each measure, scores on individual items were summed to create a composite score for each index. When necessary, questions were weighted

and/or reverse coded to make equal comparisons among responses. The scales indicated positive outcomes with higher numbers (e.g., 1=poor, 5=excellent).

Analyses were conducted in two steps. First, we generated descriptive statistics and tests of statistical significance for each index. *T*-tests (gender and representation) and analysis of variance (race/ethnicity) were performed to determine whether any differences in index means among groups were truly different or better explained as chance. Tests were performed across the following groups (equality of variance not assumed):

- Gender: Male/Female
- Underrepresentation in Computing: Underrepresented groups—Black/African-American; Latino/Hispanic; Native American and Pacific Islander; and Women. Not-underrepresented groups: Asian, White/Caucasian, Other
- Race/Ethnicity: White; Asian; Latino/Hispanic; Black/African-American; Native American and Pacific Islander; Other

Second, we used correlation and linear regression analysis to determine the relationships among the variables, with the particular goal of understanding the greatest predictors of the relationship between the eight indices in Table 1 (independent variables) and the intention to major (dependent variable).

# 5. RESULTS

## 5.1 Index Ratings and Cross-Group Comparisons

Racism/Sexism. Research shows that students exposed to stereotypes about their group can internalize and confirm them in their own behavior, leading to lower performance and departure from the major (i.e., stereotype threat) [8]. Students are also more likely to leave their major when they feel they are being treated differently as a result of belonging to an underrepresented group. The index is the sum of responses to questions about different treatment on the basis of race or ethnicity, gender, and about students telling racist or sexist jokes. Students reported few observations of seeing others treated better or worse because of race or sex or of having overheard others making sexist or racist jokes. The overall score on the scale was 29 out of 30, with 30 being very favorable and 12 or below being unfavorable. Despite the high rating, the news on racism/sexism is not all good. There was a statistically significant difference between students not underrepresented in computing and students who are underrepresented in computing, with underrepresented students perceiving slightly more racism and sexism (p=.04). There were no other differences across groups.

Meaningful/Relevant Assignments. Teaching concepts in appealing contexts and relating material to students' prior knowledge and interests is positively associated with retention. Students responded to questions about their interest in the assignments as well as the relevance of assignments to their future careers and to society. Students rated their experience in this index as "borderline favorable" (11/15, where the less than favorable range was 6 or below), indicating that while the situation is not dismal, there may be room for improvement. The low rating could be due to not enough items in the scale; more data is needed to understand students' preferences in assignments

and non-CS interests, for example. There were no statistically significant differences across groups.

Collaborative Learning. Research on collaborative learning both in general and in CS in particular show that collaborative environments lead to the development of peer networks, better self-assessments of progress, greater retention in the major, and often, improved learning. Questions in this index were focused on perceptions of working together on assignments and participating in study groups. On average, students rated collaborative learning as "borderline favorable" at 16/25 (where a very negative rating would be 10 or below and a favorable rating would be 20 or higher). There were no statistically significant differences across groups.

Teaching Assistants. TAs assigned to support labs or classes can have either positive or negative impacts on a student's experience. This index comprised questions about TAs' availability, encouragement, and the quantity and quality of help received from the TA by the student. Students rated TAs as "borderline favorable," at 21/35 (where the negative range would be 14 or below and the favorable range would be 28 or above). Members of underrepresented groups rated TAs more highly than did not underrepresented students (p=.04).

Classroom Climate. A positive classroom climate can lead to greater retention by creating natural opportunities for students to develop important peer networks, in addition to other positive conditions. This index is the sum of responses to questions about the quality of the classroom instruction, the way feedback was handled, and the level of comfort students felt in class. On average, students rated the classroom climate as "borderline favorable" at 20/30 (where the negative range is 12 or below and the minimum favorable range is 19). We found no statistically significant differences across groups.

Student-Student Interaction. Retention is expected to increase when students feel like they belong to the intellectual and social community of their major. Student-student interaction contributes to an academic community that can help overcome stereotypes and build support groups. The focus of this index was interactions among students both in and out of the classroom and lab as well as both regular school and non-school activities (e.g., watching television, exercising with students from class). Overall, respondents rated student-student interaction in the "below favorable" range at 17/30 (where a minimum of 19 is required for a favorable rating). Underrepresented students rated student-student interaction significantly better than did not underrepresented students (p=.008). This finding is contrary to what might be expected in a majority institution. Nevertheless, the mean for the underrepresented group was only "borderline favorable," at 19/30.

Pace/Workload/Programming Experience. Previous programming experience is shown to positively correlate with success in introductory programming courses. Experienced students perform better, while inexperienced students often drop out of the major. In this index, students responded to questions about the number of hours spent on homework and how well their previous experiences helped them manage the required workload. On average, students rated pace/workload as "below favorable" at 12/20 (where a minimum of 13 would be in the lower boundary of favorable and 8 or below would be less than favorable). Student members of underrepresented groups gave a significantly higher

rating than not underrepresented students (p=.02), though still below the favorable mark at 12.

Faculty-Student Interaction. Faculty-student interaction is considered another critical factor for student engagement and retention. When students are encouraged by their professors, see them as sources of advice, and feel comfortable talking to them, they are less likely to switch out of a major. Questions in this index comprised those related to advice professors give students regarding the major and career options, feedback on homework and in-class questions, and comfort talking to the professor. Respondents rated faculty-student interaction as "below favorable" at 16/35 (where the negative range is below 15 and the favorable range is 22 or above). We found no significant differences among student groups.

## 5.2 Regression: Predicting Intention to Major

We used linear regression to examine relationships between the indices and intention to major in this sample. In particular, we were interested in understanding which indices accounted for the most variance in student intention to major. We conducted regression analysis along the eight indices described above as well as gender, race/ethnicity, and representation. As can be seen in Table 3, student-student interaction, positive perceptions of pace/workload/experience, and male gender predict intention to major (variables with significant power to predict are in bold). The variable accounting for the most variance was student-student interaction. No other index was found to be a significant predictor of intention to major.

In correlation analysis, student-student interaction was positively associated with other indices, including positive perceptions of both collaborative learning ($r^2$=.546, p=.000), classroom climate/pedagogy ($r^2$=.455, p=.000), and being a member of a not-underrepresented group ($r^2$=.-254, p=.007).

## 6. DISCUSSION

This study measured eight aspects of background and in-major experience of students who took an introductory programming course at one university during the 2006-2007 academic year. It compared results for different student groups and used regression analysis to provide some understanding as to which aspects of the students' experience were predictive of intention to major.

Student-Student Interaction. Student-student interaction received a "below favorable" rating overall, yet was the strongest predictor of intention to major. This suggests that those students who were able to develop peer networks within the major were more likely to remain in the major than those who are less able. This finding, though at the department and not institution level, is consistent with Astin's Theory of Student Involvement [2], which argues that students learn more when they are more involved with both the social and the academic environment of their institution. Student-student interaction was highly correlated with positive perceptions of classroom climate and opportunities for collaborative learning. Based on the index analysis, it is not clear whether certain groups found it easier to engage in student-student interaction than others. This suggests that increasing student-student interaction must be mainstreamed, rather than being done as part of a special program for certain groups. An easy conclusion to draw is that some students are just more outgoing than others, a variable not tested in this study. However, this does not exonerate faculty wishing to increase retention in their programs from implementing classroom practices that support

student-student interaction. Special programs only work for students who participate in them. The strong relationship between collaborative environments and classroom climate suggests that faculty engineer student-student interaction by setting clear expectations for student peer involvement in their classrooms and labs through shared assignments, group problem solving, group discussions, and other methods.

**Table 3: Statistics on Intention to Major by Variable\***

| | Unstd Coefficients | | Std Coeffi- cients | | |
|---|---|---|---|---|---|
| | B | Std. Error | Beta | t | Sig. |
| (Constant) | 1.10 | 1.23 | | 0.89 | 0.37 |
| Classroom Climate/ Pedagogy | 0.03 | 0.02 | 0.16 | 1.32 | 0.19 |
| Collaborative Learning | -0.02 | 0.02 | -0.12 | -0.97 | 0.34 |
| Faculty-Student Interaction | -0.01 | 0.02 | -0.05 | -0.49 | 0.62 |
| **Student-Student Interaction\*** | **0.05** | **0.02** | **0.34** | **2.80** | **0.006** |
| Meaningful/Relevant Assignments | 0.02 | 0.033 | 0.06 | 0.68 | 0.50 |
| Teaching Assistants | -0.02 | 0.02 | -0.15 | -1.37 | 0.17 |
| Sexism/Racism | -0.03 | 0.04 | -0.07 | -0.69 | 0.49 |
| **Pace/Workload/ Experience\*** | **0.07** | **0.03** | **0.23** | **2.42** | **0.017** |
| Underrepresented | -0.08 | 0.19 | -0.05 | -0.44 | 0.66 |
| **Male gender\*** | **0.55** | **0.27** | **0.23** | **2.07** | **0.04** |
| Race/Ethnicity | 0.07 | 0.05 | 0.15 | 1.41 | 0.16 |

\*Variables in bold are three best predictors of intention to major.

Pace/Workload/Experience. Introductory classes should weed students out based on ability and potential, not on the weight of the workload or a tacit expectation that they should know more than they do. The study here shows that when students perceive the pace and workload to be too heavy in the introductory class, especially in relation to their experience level, they are unlikely to pursue the CS major. This finding is also consistent with other research in CS (see [7] for a comprehensive review of this literature). It is clear that in departments without resources to offer a "CS0" course, it is going to be more difficult to overcome the experience gap. Although it is not the intention of these authors to promote collaborative learning above other potential solutions, it is an inexpensive and doable solution (e.g., pair programming [10][11] or other in-class collaborative methods).

Male Gender. Another unsurprising finding, male gender was also positively correlated with intending to major in CS. Female students did not rate the items in the eight indices much differently than the male students; however, several important indices were below favorable, including faculty-student interaction, student-student interaction, and pace/workload/ experience. Several other indices were only borderline. Reasons for lower intention to major by women are very likely explained in the literature. For example, it is well established that women enter the computing major with less programming experience than do their male peers [7]; their participation in Advanced Placement

Computer Science has also been declining at a steeper rate than the decline of men over the past several years, down to 13% in 2007 from 20% in 2000 [12]. Several other gender differences can account for the lower intention of women to major in CS after taking the introductory course. Although research shows that the women who leave CS do so with higher grades than the men who stay [15], they have very limited tolerance for grades they perceive to be lower than those they received in high school [9]. Evidence for other differences across the genders is also strong, such as women's lower tolerance of poor teaching, limited faculty-student encouragement, and involvement with other students in the major [7]. Men, on the other hand, continue to be encouraged into majors that will make them good breadwinners (while parents place earning potential as secondary to their daughters' happiness in careers).

## 7. CONCLUSIONS

CS departments hoping to retain students can take steps to improve student-student interaction in their programs. Committees focused on undergraduate curriculum can also consider the way that students are taught, making recommendations to their colleagues that they find ways to mainstream student peer involvement in classes, such as use of collaboration for non-graded assignments and other in-class problem solving activities. By doing so, they may accomplish both improved student-student interaction and reduce the experience gap in programming among students.

## 8. ACKNOWLEDGMENTS

## 9. REFERENCES

[1] Astin, A. W. (1977). *Four critical years. Effects of college on beliefs, attitudes, and knowledge*. San Francisco: Jossey-Bass Publishers.

[2] Astin, A.W. (1984). Student Involvement: A developmental theory for higher education. *Journal of College Student Personnel*, 25, 297-308.

[3] Astin, A. W. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education.* New York: American Council on Education/Macmillan.

[4] Barker, L. J. & Garvin-Doxas, K. (2004). Making visible the behaviors that influence learning environment: A qualitative exploration of computer science classrooms. *Computer Science Education, 14*(2), 119–145.

[5] Barker, L., Cohoon, J. M., & Schaefer, S. (2007). Student Experience of the Major Survey at a Glance. http://www.ncwit.org/pdf/SEM_AtaGlance.pdf

[6] Biggers, M., Brauer, A., and Yilmaz, T. (2008). Student perceptions of computer science: A retention study comparing graduating seniors with CS leavers. *SIGCSE Bulletin ( 40)* 1, 402-406.

[7] Cohoon, J. M. & Aspray, W. F. (2006). Women and information technology: Research on underrepresentation. MIT Press: Cambridge, MA.

[8] Correll, S. (2004). Constraints into preferences: Gender, status and emerging career aspirations. *American Sociological Review (69)*1, 93-113.

[9] Jagacinski, C. M., Lebold, W. K., & Salvendy, G. (1988). Gender Differences in Persistence in Computer-Related Fields. *Journal of Educational Computing Research, 4*(2), 185-202.Kuh, G.D., Schuh, J., Whitt, E. & Associates. (1991). *Involving colleges*. San Francisco, CA: Jossey-Bass.

[10] Jacobson, N. & Schaefer, S. (2008). Pair programming in CS1: Overcoming objections to its adoption. *ACM SIGCSE Bulletin (40)*2, 93-96.

[11] McDowell, C., Werner, L., Bullock, H., & Fernald J. (2006) Pair programming improves student retention, confidence, and program quality. *Communications of the ACM (49)*8, 90-95.

[12] National Center for Women & IT. (2007). NCWIT Scorecard: A report on the status of women in information technology. www.ncwit.org/pdf/2007_Scorecard_Web.pdf

[13] Peltier, J.W., Laden, R., & Matranga, M. (1999). Student persistence in college: A review of research. *Journal of College Student Retention, 1*(4), 357-75.

[14] Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco, CA: Jossey-Bass. 50-51.

[15] Strenta, A. C., Elliott, R., Adair, R., Matier, M., & Scott, J. (1994). Choosing and Leaving Science in Highly Selective Institutions. *Research in Higher Education, 35*(5), 513-547.

[16] Tinto, V. (1994). *Leaving college: Rethinking the causes and cures of student attrition*. (2nd ed.). Chicago: The University of Chicago Press.

[17] Weidman, J. (1989). Undergraduate socialization: A conceptual approach. In J. Smart (Ed.), *Higher education: Handbook of theory and research*. New York: Agathon.