Results of a Large-Scale, Multi-Institutional Study of Undergraduate Retention in Computing

Lecia Barker
School of Information
University of Texas at Austin
Austin, TX USA
lecia@ischool.utexas.edu

Christopher Lynnly Hovey
Department of Sociology and Anthropology
Northeastern University
Boston, MA USA
c.hovey@neu.edu

Leisa D. Thompson
Department of Engineering and Society
University of Virginia
Charlottesville, VA USA
leisadthompson@gmail.com

***Abstract*—The recent upsurge in enrollments in computing means that student attrition has a substantial opportunity cost. Admitting a student who leaves both reduces graduation yield and prevents another equally qualified student from enrolling. Professors cannot change the background of students, but they can control many aspects of student experience in the computing major. This paper presents the results of a study to understand strongest predictors of retention in undergraduate computing based on a large-scale survey administered in 14 U.S. institutions. Although some factors have more influence for certain demographic groups, findings from this data set suggest that some teaching practices have more power for predicting retention in computing including: relevant and meaningful assignments, examples, and curriculum; faculty interaction with students; student collaboration on programming assignments; and for male students, pace and workload expectations relative to existing experience. Other interactions such as those that a student has with teaching assistants or peers in extracurricular activities seem to have less value for predicting retention. Faculty would be wise to protect their enrollment investments by inspecting course themes, assignments, and examples for student interest and ensuring that students have many opportunities to interact with faculty both in and outside of class.**

Keywords—student retention; persistence; undergraduate computing; assignments; women in computing; diversity in computing

I. INTRODUCTION

Retention is an important goal for colleges and universities. Each student admitted is an investment of time and resources. Students who are selected for admission also represent an opportunity cost: other students, whose pre-college abilities may be only marginally "worse," cannot be admitted because of this choice. And while college success is correlated with prior success, studies show that experiences in college are at least as important, if not more important [1], [2] suggesting that the margin of difference among students may be a poor predictor of success in and after college. Students who are admitted and then leave reduce graduation yield and waste precious space in classes and laboratories as well as the time of faculty, advisors, and teaching assistants. Studies of undergraduate retention most often focus on the likelihood that a student will remain enrolled in classes and ultimately graduate with a bachelor's degree. Such studies consider "retained" any student who graduates from the college or university, even if the student started out with the intention to major in one field, then switched to another. Studies of retention at the institution level generally focus on students' background characteristics (e.g., gender, race/ethnicity, and high school experiences); institutional factors (e.g., size, specialization, classification, and selectivity); and student integration into social and academic communities [3], [4].

Few studies of college retention focus on particular disciplines. More often, sets of disciplines are combined, especially into "STEM," or science, technology, engineering, and mathematics, such as in Seymour's and Hewitt's excellent qualitative study on why undergraduates leave the sciences [5] and in quasi-experimental studies [6], [7]. Yet the individual disciplines represented in these studies are different, as evidenced by the greater attrition of women from some fields than from others.

While there is clearly variation from one department to another based on institutional differences and individual personalities, one can argue that a disciplinary field, such as biology or computer science, constitutes a community of practice within which there are shared values for knowledge and skills, typical ways of accomplishing teaching and learning goals, and typical communication styles. Communities of practice are informal networks through which practitioners in a field come to share meaning and build knowledge [8], [9]. The apprenticeship model of doctoral education creates legacies of networks among faculty, who learn with and from each other, then go off to distant settings to become professors. Local, national, and international venues make it possible for practitioners to co-construct meaning and share practices that are common in their work. Thus it would be unsurprising to find certain similarities in terms of the ways knowledge and skills are taught and valued among departments that are geographically distant, yet distinct from other departments locally.

When studies of retention have focused on particular disciplines, they often study individual differences and pre-college experiences of students as predicting factors. Such studies typically assume that there is some deficiency among some students that needs to be "fixed" in order to make them successful, whether this is a pre-college deficiency (i.e., "mathematics readiness") [10], the need for coaching and mentoring in order to fit into or tolerate an existing culture [11], [12], or the impact of participation in groups or learning communities external to the department [13]. Generally

978-1-4799-3922-0/14/$31.00 ©2014 IEEE

excluded as predictors are the culture, social climate, and typical teaching and learning practices that students experience within a discipline. College staff and faculty cannot change students' pre-college experiences and characteristics, but they can address a student's experience of being in a major, taking classes, studying with other students, and participating in extracurricular social and educational events held by a department.

In [14], we described a model for considering systemic reform that would "change the system, not the student." The model is based in student experience of the computer science major. This model includes six interrelated components that shape students' choices to enter or remain in the major, including:

- Recruitment into early classes, but also treating early classes as recruiting opportunities, since students can still leave the major with little penalty;
- Classroom and laboratory pedagogy, which can provide a venue for the natural development of peer networks and in which students begin to be part of the community of practice;
- Curriculum, concerned with content of classes, including the degree to which assignments and examples are relevant and meaningful to students' past experiences and life goals;
- Student support mechanisms, including faculty-student interaction and student-student interaction, giving rise to a sense of belonging in a community;
- Institutional policies related to admissions, faculty roles and promotion, and class taking requirements; and
- Evaluation and tracking of student outcomes, providing data for assessing whether students are learning and whether some students are better served than others.

For fields with the worst record for gender parity, such as computer science, it is important to understand why some students stay and some students leave. Women in computing fields leave at a higher rate than men [15], despite having higher grades than the men who stay [16]–[19]. The study presented below focuses on the factors of retention that faculty and academic departments can modify to protect their enrollment investments (i.e., to retain students). Below, we present a large-scale, survey-based study of retention factors to explore the biggest predictors of a computer science major's intention to complete the major. Our retention model, based on the student's experience of the computer science major as discussed above, includes primarily factors faculty have the power to modify. Faculty are committed to retaining students, but often have little time or resources for doing so. Therefore, it is useful to understand which retention factors will bring the greatest return on investment. Therefore, we use our survey data to address these questions:

Among factors that faculty can control, which are the strongest predictors of retention?

How do predictors of retention vary across student groups?

First, we describe our choices for constructing the "Student Experience of the Major" survey. Next, we describe the sample profile and results for predicting retention in computer science. We conclude with a discussion of the findings.

II. FACTORS RELATED TO RETENTION

The Student Experience of the Major (SEM) survey is based in the reform model summarized above and in [14] and can be viewed in [20]. That model is based partly on general scholarship related to student engagement and retention (summarized in [21]–[23]). This scholarship views retention and voluntary attrition as intricately linked to a longitudinal process in which individual students assign meaning to formal and informal interactions during their educational experience. As a consequence, students with higher levels of social and academic integration will be more likely to persist in their major. Integration is accomplished through experience of the factors we list below. In addition, the SEM survey takes into account research specific to computer science education. The major factors contributing to retention are explained below. All but the last, commitment to the major, are independent (predictor) variables. Commitment to the Major is the dependent variable.

A. Classroom Climate

Learning theorists argue that learning occurs as a situated accomplishment within the constraints of the everyday practices of a social environment [24], [25] and mediated by communication typical to those environments [26]–[28]. One of the implications of this view is that when one learns, one develops a new or "different sense of self as a consequence of active participation in a social practice" [29]. Classrooms are the predominant venue in which students are exposed to the skills and knowledge, knowledge presentation, and expectations about the kinds of people who belong (or not) in a degree program. How faculty talk about students in class helps to cement students' views of whether they belong (or not). For example, instructors who unwittingly align being "smart" with having prior programming experience can negatively impact the perceptions of students with less experience [30]. Students who "posture" or "show off" in class can intimidate other students [30]–[32]. Feeling that one doesn't belong, combined with being a minority in the classroom and confronting various gendered stereotypes, results in a loss of confidence for women, which can then lead to switching major [31], [33]. In the SEM survey, we created a composite variable by combining ratings of teaching quality with ratings of classroom climate. When we conducted reliability tests to determine if these individual items should be combined into one composite scale, we found that the items were not closely enough correlated for us to feel confident to combine them. Thus, the variable for classroom climate omits ratings of teaching quality and only encapsulates students' perceptions about comfort and frequency with asking questions in class, and how frequently faculty call on students by name in class.

B. Prior Experience and Workload

Students want to be challenged in class, but not beyond their ability or time to absorb content and complete work. Retention within a major is predicated on students believing that their academic efforts are a good use of their time [23].

However, prior experience with programming is positively related to success in introductory computing courses [16], suggesting that instruction is not appropriate for a beginner. The composite variable for Prior Experience and Workload captures students' responses to questions about the number of hours spent on homework, pace of classes, and how well their previous exposure and experiences within computing have helped them manage the required workload.

C. Relevant, Meaningful Assignments

Studies in computer science show that the types and topics of assignments, examples, and classes included in introductory courses appeal differently to different groups of students [31], [34], [35]. Computer science is an abstract discipline; as a result, lectures can focus on abstract principles, separated from the world of real-life application [34], [36]. The degree to which faculty are able to relate to students' past experience and knowledge, values, and future goals is related to retention [23]. In our survey, we created a composite variable called Relevant and Meaningful Assignments, in which students indicated how many assignments were interesting to them and rated their interest in and understanding of personal relevance of assignments.

D. Student-Faculty Interaction

Student-faculty interaction is often considered the largest contributor to retention in engineering and computing [3]. Student-faculty interaction occurs both in and outside of class, but informal interaction that extends beyond formal contact is shown to have several strong benefits [37], [38]. Students are more likely to attend classes in which faculty show enthusiasm for the material, who are available for questions after class or during office hours, and who develop a reputation for being personal and approachable [39]. These findings are especially strong for women students [5], [39], [40]. In the SEM survey, questions representing this factor reflect student perception of professors' advice regarding the major and career options, feedback on homework and in-class questions, faculty encouragement, and comfort talking to the professor.

E. Student-Teaching Assistant Interaction

A few studies have examined the quality of interactions with teaching assistants (TAs) (e.g.,[41], [42]), though few have specifically related student-TA interaction to retention. One large-scale study found that student interaction with teaching assistants can influence retention by affecting lab climate, awareness of course grades, and students' knowledge of careers [43]. A challenge in studying the impact of TAs is the wide variation in how faculty utilize teaching assistants. The SEM survey operationalizes Student-Teaching Assistant Interaction with questions about TAs' availability, encouragement, and the quantity and quality of help received from the TA by the student.

Table 1: SEM Factors and Corresponding Survey Topics

Dimension	Survey Measures
Commitment to the Major	Likelihood of declaring or completing a computing major or minor; sense of belonging in the computing major
Classroom Climate	Perceived quality of classroom instruction; speed of feedback to students on assignments; professors' use of student names; level of comfort in class
Prior Experience & Workload	Hours spent on homework and labs; pace of classes; perception of experience required for assignments compared to students' experience levels
Meaningful Assignments	Interest in course assignments; relevance of assignments to personal interests, society, and career; understanding the relevance of content to career
Student-Faculty Interaction	Career and academic advice from professors; professor praise and encouragement; professor mentoring; comfort in talking to the professor one on one
Student-TA Interaction	Frequency of and comfort with asking TAs for help; rating of TAs' abilities to support learning and provide encouragement
Within-Major Social Life	Spending non-school time with other students from class; encouragement from other students to persist; membership in student organizations
Collaboration on Programming Assignments	Encouraged by faculty to collaborate on programming assignments
Egalitarian Environment	Belief that some students are treated differently by race/ethnicity or gender; prevalence of racist/sexist jokes in class or lab

F. Within-Major Social Life

Extracurricular student interaction with peers with similar interests is thought to increase retention, but the type of interactions that retain students is understudied [44]. For example, "learning communities" of students who have similar or the same major have been established in dormitories so that these students are frequently in contact. The theory is that one can develop social networks that help ease the burdens and responsibilities associated with the major. Students exchange information that can enhance learning experiences and beliefs about belonging [45]. The SEM factor for Within-Major Social Life asked about frequency and quality of student interactions during non-school time and in student groups (e.g., ACM student groups).

G. Collaborative Assignments

Students working together have more opportunities to determine who is knowledgeable and who is not, also, possibly overcoming stereotypes about certain types of people. Collaborative learning can facilitate student-student interactions and peer support, providing opportunities for students to articulate what they are learning and hear their peers do the same; these practices are shown to improve learning and sense of belonging [46]–[48]. Studies have also shown that collaborative assignments can increase motivation, help students recognize how to apply unique skills to group efforts, increase awareness of project goals, and share knowledge beyond that required to fulfill the immediate assignment [49]. In computer science, pair programming has been associated with taking additional classes [50]. The SEM survey includes several questions that attempt to capture faculty attitudes toward allowing students to work together in contrast to beliefs that working together could be perceived as cheating. It also included the degree to which students join study groups. This composite variable failed reliability tests. Therefore, in the regression model, we used an instrumental variable describing the frequency with which students are encouraged to Collaborate on Programming Assignments.

H. *Egalitarianism*

Given the male-dominated nature of computing disciplines, we included items on sexism and racism to understand how they might affect retention in computer science. Students are less likely to stay in a major when they feel they are being treated differently as a result of belonging to an underrepresented group. Repeated, subtle messages that one is not like the other students—not as smart, not interested in the same activities, not a "real" computing person—make it difficult to imagine oneself developing that identity. In addition, students may internalize or seek confirmation in their own behavior when exposed to stereotypes about their gender or racial group (i.e., stereotype threat). Experimental studies show that stereotype threat can reduce performance and interest and change career choices [51], [52]. Asking questions about sexism and racism is tricky, because most people do not want to perceive themselves either as a perpetrator or a victim. Also, more concerning today are the implicit biases that shape perceptions and choices and mark people as different in very subtle ways. Asking people if they think they were discriminated against does not necessarily mean they were or were not, since what happened to leave that impression can be subtle or difficult to prove. The SEM factor representing Egalitarian Environment combines responses to questions about observations of differential treatment of others (not of oneself) based on race/ethnicity or gender, as well as observations of other students telling sexist or racist jokes.

I. *Commitment to the Major*

Though there are competing definitions of attrition and retention, for the purposes of this study we are concerned with the self-reported likelihood of declaring or completing a computing major or minor [53]. We use Commitment to the Major as a proxy for retention; it is the main dependent variable used in this study. Commitment to the Major combined three categories of questions, including likelihood to complete the major, likelihood to pursue a career in the field, feelings of belonging in the major.

III. PROCEDURES, SAMPLE, AND ANALYTICAL METHODS

A. *Survey Development and Procedures*

The SEM survey was developed and piloted in 2006 and revised for administration between 2008 and 2011. The original pilot data are not included in this analysis. Surveys were administered online using SurveyMonkey. In an effort to reduce selection bias, each participating department determined the incentives that were offered in exchange for students' participation. Relevant authorities, including the Institutional Review Board for each school, oversaw that research conformed to ethical standards. Students responded anonymously to a set of questions standardized across the institutions concerning their experiences in and perceptions of the major. Students were also asked to report on their majors and reasons for declaring (or not declaring) a computing major, as well as the measures concerning their intention to persist.

B. *Student Sample Profile*

In total, 2,077 students from 14 institutions responded in part to the survey, with 1,312 students replying to all questions used in the regression analysis. (The high level of missing data is primarily due to some participating institutions' decisions to exclude certain questions.) Twenty percent of respondents were female (n=1,945), and 8.1 percent of students belonged to a group that is a racial minority underrepresented in computing (n=1,910). Whites and Asians are not considered underrepresented. In total, 25.8 percent of students reported being a minority in computing by either gender or race (n=1,892); slightly fewer than 2 percent reported being a minority by both gender and race (n=1,892). Students were on average 23 years old ($S = 4.172$, $n = 1566$) and were evenly distributed by academic level, with between 19.8 and 27.1 percent falling into each of the four years of college. Roughly eight percent of students reported being "fifth year seniors."

C. *Institutional Differences in Sample*

On average, each of the 14 institutions had roughly 148 respondents, with a range between 44 and 451 students ($S = 115.968$, median = 104.5). Significant differences between institutions existed in gender representation ($\varphi = 0.245$, $p<.001$) and racial minority representation ($\varphi = 0.195$, $p<.001$), with student compositions ranging from 6.4 to 59.3 percent female and from 0.0 to 20.6 percent racial minority. Differences were also found in composition of academic year. The distribution of academic year by institution was statistically significant, ($\chi2(df=52) = -0.106$, $p<.001$) due to seven schools with a concentration of students at least one standard deviation above or below the average in a specific year for all schools. We found no difference in academic year between genders ($t = 1.390$, $p=0.165$). However, there was a difference in academic year between racial groups ($\chi2(df=12) = 33.894$, $p<.001$): White students were less likely to be freshmen and more likely to be upper classmen, while Asian students were less likely to be juniors and more likely to be freshmen or sophomores. Underrepresented and mixed racial groups were far more likely to be freshmen, and least likely to be any other year. We do not know how to explain this: it may suggest greater attrition among underrepresented students early in the major or recent, successful recruitment efforts.

D. *Analytical Methods*

As described in Section II, we developed ten composite variables from a total of 70 survey items. For each composite variable, individual item scores were summed to create a composite score; when necessary, questions were weighted and/or reverse coded to make equal comparisons among responses. The internal consistency of composite variables was tested using Cronbach's Alpha. As discussed in Section II, all but two of our factors ("classroom climate and pedagogy" and "collaborative learning") met a minimum standard of $\alpha \geqslant$

0.6. We also employed factor analysis to confirm the composite variables' reliability using principal components extraction with varimax rotation. The results confirmed the reliability tests using Cronbach's Alpha. We concluded that 1) classroom climate and teaching were not related enough to justify combining them; 2) our operationalizations of classroom climate and collaborative learning questions were faulty. For the classroom climate/teaching variable, we parsed down the constituent items to become more reliable. In the case of collaborative learning, we selected one instrumental variable from the survey that best operationalized the concept, which although it came out as a predictor in most of the models, may have reduced its predictive power.

E. Explanation of the Regression Models

We used the independent composite variables to develop and test seven models to understand the strongest predictors of "Commitment to the Major" taking into account different groups of students, as explained below. Each model affords a different way of looking at the data. All variables in the regression models passed minimum standards for variance inflation factors and tolerance measures, indicating that no variables were risks for multicollinearity. Six models used ordinary least squares regression. The final model used mixed linear effects in order to determine whether there were systematic biases based on institutional differences. The outcomes of the analyses are shown in Tables 2 and 3. Because these models use factored variables, the interpretation of coefficients is not straightforward. However, these models allow us to compare the relative effects of each because they used standardized scales. Independent variables that were the strongest predictors of the outcome variable, Commitment to the Major, are shown in **bold** within cells. The level of significance (the chance of the findings occurring randomly) is indicated with stars below the tables. A cross indicates borderline cases (p<.075).

<u>All Students: All-1 and All-2</u>. The two models shown in Table 2 both use ordinary least squares regression and include the entire data set, hence their names All-1 and All-2. All-1 controls for academic year and age. In other words, academic year and age are held constant to better understand the influence of the other variables. Although academic year and age were positively correlated ($r = 0.641$, $p<.001$), each passed variance inflation factor and tolerance tests in the regression models, meaning that multicollinearity is not a risk. All-2 is similar to All-1, but adds an interaction variable for being *both* female and a racial minority member.

<u>Gender and Race/Ethnicity</u>. To identify the degree to which different variables predicted intention to major in homogeneous groups, the next four models isolated subsets of the population based on gender and race. This is shown through shading in Table 3: under the columns Female, Male, Under-rep, and Majority are the results of ordinary least squares regression for these groups independent of the other groups. Isolating demographic groups and running separate ordinary least squares models better illustrate the relative importance of the independent variables to each group.

<u>Controlling for Institutional Influences</u>. The final model was a linear mixed effects model that controlled for "Institution," which helps to identify unmeasured effects that may stem from local departmental differences. This model helps to address the question of whether departmental, institutional, and peer cultures locally affect students' intentions to complete a major [54]. For example, in an institution where students are more likely to work to earn income or take out more student loans, retention may be lower because computer science majors are attracted to the income possibilities, with or without the degree. A variable for institution was used as a fixed effect and the other composite variables included as random effects. This approach allowed us to control for possible unmeasured variables that are common to all students in an institution that might otherwise bias the results. This model is represented in Table 3 under the right-most column and titled "Inst Diffs."

IV. Results of the Seven Models

A. All Students Models, Controlling for Age and Level

These two models each include the entire data set, meaning that all students are treated as one group, but controlling for academic year and age. When the students are treated as one group, statistically significant and positive predictors for the dependent variable Commitment to the Major include: Meaningful Assignments, Student-Faculty Interaction, Male Gender, Collaboration on Programming Assignments, and Prior Experience and Workload in order of strength of prediction. Classroom Climate was a negative predictor in both models.

Several variables had no significant influence on Commitment to the Major, including Race (Underrepresented Minority and Majority), the interaction of race and gender, academic year, and age. This finding challenges other studies that suggest that the first two years of a students' college experience are when students are most at risk for dropping out of the major [55]. In addition, student-teaching assistant interaction, within-major social life, and egalitarian environment were not significant predictors of retention.

Table 2: Ordinary Least Squares Models 1 and 2

Predictor Variables	All-1	All-2
Female	**-0.275*** (.057)**	
Racial Minority	-0.182 (.094)	
Gender *and* Racial Minority	-0.028 (.197)	
Academic Year	-0.013 (.021)	.000 (.020)
Age	-0.005 (.006)	-0.007 (.006)
Classroom Climate	**-0.123* (.051)**	**-0.116* (.050)**
Prior Experience/Wrkload	**0.081* (.032)**	**0.096** (.032)**
Meaningful Assignments	**0.378*** (.043)**	**0.393*** (.044)**
Student-Faculty Interaction	**0.266*** (.053)**	**0.279*** (.053)**
Student-TA Interaction	-0.004 (.034)	-0.032 (.033)
Within-Major Social Life	0.001 (.031)	-0.007 (.031)
Collaboration	**0.108*** (.023)**	**0.110*** (.023)**
Egalitarian Environment	-0.002 (.049)	0.028 (.048)
Intercept	1.461*** (.270)	1.236*** (.266)
Adjusted R-Squared	0.163	0.136

Standard error in parentheses. ***p<.001; **p<.01; *p<.05

Table 3: Female, Male, Under-represented minority, Majority, and Linear Mixed Effects Models

Predictor Variables	Female	Male	Under-rep	Majority	Inst Diffs
Classroom Climate	-0.268†	-0.082	-0.153	-0.106*	-0.019
	(.138)	(.047)	(.154)	(.049)	(.045)
Prior Experience & Workload	0.059	**0.121*****	**0.255***	**0.099****	**0.132*****
	(.076)	(.030)	(.113)	(.030)	(.028)
Meaningful Assignments	**0.579*****	**0.372*****	**0.469****	**0.408*****	**0.434*****
	(.123)	(.039)	(.138)	(.041)	(.037)
Student-Faculty Interaction	**0.295***	**0.258*****	0.199	**0.253*****	**0.379*****
	(.130)	(.049)	(.164)	(.051)	(.048)
Student-TA Interaction	0.117	-0.062†	-0.245	-0.029	**-0.118*****
	(.079)	(.031)	(.133)	(.031)	(.030)
Within-Major Social Life	0.003	0.015	0.004	-0.005	(.041)
	(.075)	(.030)	(.120)	(.030)	(.028)
Collaboration	**0.157****	**0.091*****	0.136	**0.106*****	**0.058****
	(.058)	(.021)	(.078)	(.022)	(.022)
Egalitarian Environment	-0.108	0.063	0.105	0.021	0.076
	(.107)	(.046)	(.151)	(.046)	(.041)
Intercept	0.562	1.031***	0.785	1.079***	0.320
	(.553)	(.222)	(.815)	(.226)	(.231)
Adjusted R-Squared	0.124	0.15	0.136	0.134	

Standard error in parentheses. $^{\dagger}p<.075$; Inst Diffs uses Linear Mixed Effects to controls for institutional differences across the fourteen institutions.

B. Gender and Race/Ethnicity Models

Female Students Only. When looking at females as an independent group, the significant factors predicting Commitment to the Major were Meaningful Assignments, Student-Faculty Interaction, and Collaboration on Programming Assignments, listed in order of strength. Interestingly, Meaningful Assignments was the strongest predictor of any variable in any of the models. This is a strong recommendation for retaining women: assignments, examples, course themes should be meaningful to women's interests. This is consistent with previous qualitative studies as discussed in Section II. It is not surprising that Prior Experience and Workload was not a predictor for women, since they tend to enter computing majors with less prior experience than their male counterparts. Although sexism and racism linger in both the professional and academic fields of computing, we found no evidence in our study that sexism or racism are affecting student retention. Sadly, it is not implausible to imagine that the women and minorities who are retained (and therefore responded to the survey) are more resilient toward sexism and racism.

Male Students Only. Male students were not much different from female students. The only difference was that Prior Experience and Workload was an additional predictor. The other significant predictors of Commitment to the Major were Meaningful Assignments, Student-Faculty Interaction, and Collaboration on Programming Assignments. Meaningful Assignments is the most important predictor for both men and women, though somewhat less important for men. However, the strength of Student-Faculty Interaction is only slightly different. Collaborating on Programming Assignments was also less important for men than for women, but was still a significant predictor.

Underrepresented Minority Students Only. Meaningful Assignments came in as the strongest predictor of Commitment to the Major for minority students also, followed by Prior Experience and Workload. Given the increase in effect size in the underrepresented minority-only model as compared to other groups, we conclude that 'time spent studying', and 'matching levels of experience in the class to previous experience in computing before entering college' may be particularly important for students of color in positively influencing retention. Similar to the women-only group, it appears that increasing the relevance of curriculum, assignments, and examples could be an effective way to retain under-represented students.

Majority Students Only. For the group of students who were in the racial majority, again the strongest predictor of Commitment to the Major was Meaningful Assignments. Runners up included Student-Faculty Interaction, Collaboration on Programming Assignments, and Prior Experience and Workload in order of strength (though Collaboration was only slightly stronger than Prior Experience). Interestingly, Classroom Climate had a slight negative effect on Commitment to the Major. Although we include this result here, we have limited confidence in it due to the construction of the variable. This variable centered on how comfortable students felt asking questions, the frequency of asking questions, and how frequently professors call on students using their names. Given these results, we could tentatively conclude that majority students who feel the classroom is a comfortable climate are more likely to leave the major. Future research should investigate this in greater depth, considering its counter-theoretical nature.

Institutional Influences. Results for this model were similar to the six other models, in that Meaningful Assignments, Student-Faculty Interaction, Prior Experience and Workload, and Collaboration on Programming Assignments were statistically significant predictors of Commitment to the Major. Some differences were observed, including relatively large increases in explanatory power for Prior Experience and Workload, and for Student-Faculty Interaction, and a decrease in power by almost half for Collaboration on Programming Assignments. The Institutional Influences model had one quite different (and unexplainable) finding: Student-Teaching Assistants Interaction showed up as a slight negative predictor of retention. One possible explanation is that because students often see teaching assistants as stand-ins for faculty [39], [42], this negative effect could be due to disaffected students accessing teaching assistants in lieu of disinterested or unapproachable faculty in one or more of the campuses.

V. Discussion and Conclusions

No matter how we sliced the data, Meaningful Assignments maintained the strongest effect size across all groups and was highly significant as a predictor of Commitment to the Major in each model. Of note, the effect size substantially increased for females and underrepresented minority students as compared to males. From this evidence, we conclude that female students' commitment to computer science is more strongly predicated on completing assignments they find interesting and that place a high level of emphasis on demonstrating how useful computing can be in relation to their career goals and to society. This finding, that Meaningful Assignments is a strong predictor of women's retention, should guide educators interested in attracting and retaining more females to their computing departments while ensuring the continued commitment from majority groups. However, it is important to repeat that this factor was the strongest predictor of retention for all student groups.

Student-Faculty Interaction was also a significant predictor of retention. Partially confirming previous studies [3], [23], [56]–[58], Student-Faculty Interaction maintained a statistically significant and relatively strong, positive effect size across groups except for racial minorities. We are not sure how to interpret this finding. It is possible that minority students are less comfortable talking to professors or that professors give them less advice and encouragement. More in-depth probing of this finding in future studies may solve this riddle.

Similarly, Collaboration on Programming Assignments was moderately important and statistically significant to all groups except underrepresented minorities. For minorities, there was still a positive effect, but the large standard error prevented it from being statistically significant. A *t*-test revealed significant differences between racial minority and majority groups ($t = -2.626$, $p<.01$) with majority groups giving a higher overall perception that professors encourage students to collaborate on assignments for the class. Though there is not enough information here to explain this difference, we speculate that because majority groups perceive more encouragement to collaborate, it becomes an important factor majority students consider when staying in computer science. On the other hand, because minority students in this data set experienced less encouragement to collaborate, other factors are larger contributors for retention.

With respect to differences across departments surveyed, we found that interaction with teaching assistants was negatively associated with retention. We speculate that the reason for this may lie in differences among institutions in terms of the effectiveness of how teaching assistants are trained or used. Few departments offer training programs for teaching assistants [43], so even though students access them, they may not be able to effectively help the students, contributing to overall lower views of the major. Future research may probe deeper into this finding.

While all the factors identified in the literature review and in this study are important to improving students' experience in college, we found that the most important factor in retaining students–consistent across all groups and models–was the use of relevant, meaningful assignments. The second biggest predictor was faculty-student interaction. Departments hoping to retain the students they admit would be well advised to inspect course themes, assignments, and examples used in lectures for the degree to which they are of personal interest to students, and relevant to students' career goals and interests in society. They can also improve retention by ensuring that students have ample opportunities to interact with faculty both in and out of the classroom.

Acknowledgments

This research was sponsored by the National Science Foundation Division of Human Resource Development award #0533565. Any opinions, findings, or conclusions, or recommendations expressed are those of the authors and do not reflect those of NSF. We thank Joanne McGrath Cohoon, Matt Hersh, Anita Iannucci, Kim Kalahar, Suzanne Schaefer, Emmy Soyka, and Tim Weston for support in survey design, administration, and data analysis. Special thanks to all the computer science students who agreed to take the survey and to the faculty who made it happen on their campuses.

References

[1] G. D. Kuh, T. Cruce, R. Shoup, J. Kinzie, and R. M. Gonyea, "Unmasking the Effects of Student Engagement on College Grades and Persistence," *J. High. Educ.*, vol. 79, no. 5, pp. 540–563, 2008.

[2] G. D. Kuh, J. Kinzie, J. H. Schuh, and E. J. Whitt, *Student Success in College: Creating Conditions That Matter*. John Wiley & Sons, 2010.

[3] A. W. Astin, "Student involvement: A developmental theory for higher education," *J. Coll. Stud. Pers.*, vol. 25, no. 4, pp. 297–308, 1984.

[4] A. Seidman, Ed., *College student retention: Formula for Student Success*, 2nd ed. Lanham, MD: Rowman & Littlefield Publishers, 2012.

[5] E. Seymour and N. Hewitt, *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press, 1997.

[6] B. LeBeau, M. Harwell, D. Monson, D. Dupuis, A. Medhanie, and T. R. Post, "Student and high-school characteristics related to completing a science, technology, engineering or mathematics (STEM) major in college," *Res. Sci. Technol. Educ.*, vol. 30, no. 1, pp. 17–28, 2012.

[7] P. A. Daempfle, "An analysis of the high attrition rates among first year college science, math, and engineering majors," *J. Coll. Stud. Retent. Res. Theory Pract.*, vol. 5, no. 1, pp. 37–52, Jan. 2003.

[8] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation*, 1st ed. Cambridge University Press, 1991.

[9] Brown, A. Collins, and P. Duguid, "Situated Cognition and the Culture of Learning," *Educ. Res.*, vol. 18, no. 1, pp. 32–42, Jan. 1989.

[10] L. Moses, C. Hall, K. Wuensch, K. De Urquidi, P. Kauffmann, W. Swart, S. Duncan, and G. Dixon, "Are Math Readiness and Personality Predictive of First-Year Retention in Engineering?," *J. Psychol.*, vol. 145, no. 3, pp. 229–245, 2011.

[11] J. M. Karlen, "Attrition of women business majors in an urban community college," *J. Coll. Stud. Retent. Res. Theory Pract.*, vol. 5, no. 1, pp. 1–9, Jan. 2003.

[12] L. Wells-Glover and E. Newton, "Mentors for Undergraduates in Technical Disciplines: A Collaborative Effort by Faculty, Student Development Professionals, and Alumni to Improve Undergraduate Retention and Success in Technical Majors," *J. Coll. Stud. Retent. Res. Theory Pract.*, vol. 1, no. 4, pp. 311–321, Feb. 2000.

[13] S. Baker and N. Pomerantz, "Impact of Learning Communities on Retention at a Metropolitan University," *J. Coll. Stud. Retent.*, vol. 2, no. 2, pp. 115–26, Jan. 2001.

[14] L. J. Barker, J. M. Cohoon, and L. D. Thompson, "Work in progress — A practical model for achieving gender parity in undergraduate computing: Change the system, not the student," in *2010 IEEE Frontiers in Education Conference (FIE)*, 2010, pp. S1H–1–S1H–2.

[15] J. M. Cohoon, J. Cohoon, and S. Turner, "Departmental factors in gendered attrition from undergraduate IT majors," 2001.

[16] J. M. Cohoon and W. Aspray, "A critical review of the research on women's participation in postsecondary computing education," in *Women and Information Technology: Research on Underrepresentation*, Cambridge, MA: MIT Press, 2006, pp. 137–180.

[17] S. Katz, D. Allbritton, J. Aronis, C. Wilson, and M. L. Soffa, "Gender, achievement, and persistence in an undergraduate computer science program," *DATA BASE Adv. Inf. Syst.*, vol. 37, no. 4, pp. 42–57, 2006.

[18] M. R. H. Roberts, T. McGill, and T. Koppi, “What Students are Telling us about Why They Left Their ICT Course,” *Innov. Teach. Learn. Inf. Comput. Sci.*, vol. 10, no. 3, pp. 68–83, Nov. 2011.

[19] A. C. Strenta, R. Elliott, R. Adair, M. Matier, and J. Scott, “Choosing and Leaving Science in Highly Selective Institutions,” *Res High Educ*, vol. 35, no. 5, pp. 513–547, 1994.

[20] L. Barker, J. M. Cohoon, S. Schaefer, and J. Krauss, “Survey-in-a-Box: Student Experience of the Major,” *National Center for Women and Information Technology*, 15-May-2007. [Online]. Available: http://www.ncwit.org/resources/survey-box-student-experience-major. [Accessed: 21-Apr-2014].

[21] V. Tinto, *Completing college: Rethinking institutional action*. Chicago, IL: The University of Chicago Press, 2012.

[22] J. M. Braxton, W. R. Doyle, H. V. Hartley, A. S. Hirschy, W. A. Jones, and M. K. McLendon, *Rethinking College Student Retention*, 1st ed. San Francisco, CA: Wiley, 2014.

[23] M. Biggers, A. Brauer, and T. Yilmaz, “Student Perceptions of Computer Science: A Retention Study Comparing Graduating Seniors with Cs Leavers,” in *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education*, New York, NY, USA, 2008, pp. 402–406.

[24] J. Lave, “The culture of acquisition and the practice of learning,” Institute for Research on Learning, Palo Alto, CA, 1990.

[25] J. Nespor, “Curriculum and conversions of capital in the acquisition of disciplinary knowledge,” *J. Curric. Stud.*, vol. 22, no. 3, pp. 217–232, 1990.

[26] E. Jadallah, “Constructivist learning experiences for social studies and education,” *Soc. Stud. Sci.*, vol. 91, no. 5, pp. 221–226, 2000.

[27] N. Mercer, *Words and minds: How we use language to think together*. London: Routledge, 2000.

[28] G. Wells, *Dialogic enquiry: Towards a sociocultural practice and theory of education*. Cambridge, MA: Cambridge University Press, 1999.

[29] M. A. Eisenhart and E. Finkel, *Women's science: Learning and succeeding from the margins*. Chicago: University of Chicago Press, 1998.

[30] L. J. Barker and K. Garvin-Doxas, “Making visible the behaviors that influence learning environment: A qualitative exploration of computer science classrooms,” *Comput. Sci. Educ.*, vol. 14, no. 2, pp. 119–146, 2004.

[31] J. Margolis and A. Fisher, *Unlocking the Clubhouse: Women in Computing*. Cambridge, MA: MIT Press, 2002.

[32] L. J. Barker, M. O'Neill, and N. Kazim, “Framing Classroom Climate for Student Learning and Retention in Computer Science,” in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 2014, pp. 319–324.

[33] J. M. Cohoon, “Recruiting and retaining women in undergraduate computing majors,” *SIGSCSE Bull.*, vol. 34, no. 2, pp. 48–52, 2002.

[34] B. C. Wilson, “A study of factors promoting success in computer science including gender differences,” *Comput. Sci. Educ.*, vol. 12, no. 1–2, pp. 141–164, 2002.

[35] A. Forte, “Programming for communication: Overcoming motivational barriers to computation for all,” in *Human Centric Computing Languages and Environments, 2003*, 2003, pp. 285–286.

[36] A. Revell and E. Wainwright, “What Makes Lectures ‘Unmissable’? Insights into Teaching Excellence and Active Learning,” *J. Geogr. High. Educ.*, vol. 33, no. 2, pp. 209–223, May 2009.

[37] J. J. Endo and R. L. Harpel, “The Effect of Student-Faculty Interaction on Students Educational Outcomes,” *Res. High. Educ.*, vol. 16, no. 2, pp. 115–138, Jan. 1982.

[38] E. T. Pascarella and P. T. Terenzini, “Student-Faculty and Student-Peer Relationships as Mediators of the Structural Effects of Undergraduate Residence Arrangement,” *J. Educ. Res.*, vol. 73, no. 6, pp. 344–353, Jul. 1980.

[39] S. R. Cotten and B. Wilson, “Student-Faculty Interactions: Dynamics and Determinants,” *High. Educ.*, vol. 51, no. 4, pp. 487–519, Jun. 2006.

[40] G. Escobedo, “A Retention/Persistence Intervention Model: Improving Success Across Cultures,” *J. Dev. Educ.*, vol. 31, no. 1, pp. 12–37, Fall 2007.

[41] P. E. Dickson, “Using Undergraduate Teaching Assistants in a Small College Environment,” in *Proceedings of the 42Nd ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 2011, pp. 75–80.

[42] D. M. Shannon, D. J. Twale, and M. S. Moore, “TA teaching effectiveness,” *J. High. Educ.*, vol. 69, no. 4, p. 440, Jul. 1998.

[43] C. Cook, C. O'Neal, T. Perorazio, J. Purkiss, and M. Wright, “The impact of teaching assistants on student retention in the sciences: lessons for TA training,” *J. Coll. Sci. Teach.*, vol. 36, no. 5, p. 24, 2007.

[44] M. Andrade, “Learning Communities: Examining Positive Outcomes,” *J. Coll. Stud. Retent. Res. Theory Pract.*, vol. 9, no. 1, pp. 1–20, Jul. 2007.

[45] L. E. Anderson and J. Carta-Falsa, “Factors That Make Faculty and Student Relationships Effective,” *Coll. Teach.*, vol. 50, no. 4, pp. 134–138, Oct. 2002.

[46] E. Forman and C. Cazden, “Exploring Vygotskian perspectives in education: The cognitive value of peer interaction,” in *Culture, communication, and cognition: Vygotskian perspectives*, Cambridge, MA: Cambridge University Press, 1985.

[47] A. C. Graesser, N. Person, and J. P. Magliano, “Collaborative dialogue patterns in naturalistic one-on-one tutoring,” *Appl. Cogn. Psychol.*, vol. 9, pp. 495–522, 1995.

[48] N. M. Webb and A. S. Palinscar, “Group processes in the classroom,” in *Handbook of Educational Psychology*, New York: Macmillan, 1996, pp. 841–873.

[49] N. Nitta, T. Yasuhiro, and K. Izuru, “A practice of collaborative project-based learning for mutual edification between programming skill and artistic craftsmanship,” in *39th ASEE/IEEE Frontiers in Education (FIE) Conference, 2009*, 2009, pp. 1–5.

[50] C. McDowell, L. Werner, H. E. Bullock, and J. Fernald, “Pair programming improves student retention, confidence, and program quality,” *Commun. ACM*, vol. 49, no. 8, pp. 90–95, 2006.

[51] J. Aronson, “The threat of stereotype,” *Educ. Leadersh.*, vol. 62, no. 4, pp. 14–19, 2004.

[52] Correll S.J., “Contstraints into Preferences: Gender, Status and Emerging Career Aspirations,” *Am. Sociol. Rev.*, vol. 69, pp. 93–113, Feb. 2004.

[53] C. Campbell and J. Mislevy, “Student Perceptions Matter: Early Signs of Undergraduate Student Retention/Attrition,” *J. Coll. Stud. Retent. Res. Theory Pract.*, vol. 14, no. 4, pp. 467–493, Jan. 2012.

[54] L. Oseguera and B. S. Rhee, “The Influence of Institutional Retention Climates on Student Persistence to Degree Completion: A Multilevel Approach,” *Res. High. Educ.*, vol. 50, no. 6, pp. 546–569, Sep. 2009.

[55] M. J. Clancy, R. E. Pattis, and M. Stehlik, “Approaches to Programming Assignments in CS 1 and CS 2,” in *Proceedings of the Twenty-fourth SIGCSE Technical Symposium on Computer Science Education*, New York, NY, USA, 1993, p. 308–.

[56] S. Devadoss and J. Foltz, “Factors influencing student class attendance and performance,” *Int. Adv. Econ. Res.*, vol. 2, no. 2, pp. 194–195, May 1996.

[57] G. Clark, N. Gill, W. Marion, and R. Whittle, “Attendance and Performance: Correlations and Motives in Lecture-Based Modules,” *J. Geogr. High. Educ.*, vol. 35, no. 2, pp. 199–215, May 2011.

[58] E. Cohn and E. Johnson, “Class Attendance and Performance in Principles of Economics,” *Educ. Econ.*, vol. 14, no. 2, pp. 211–233, Jun. 2006.