A mixed-methodology study of the origins and impacts of the gender gap in college physics
Lauren E. Kost

This research study is part of a broader effort to investigate the “gender gap” documented in university physics courses. This gap manifests itself in a variety of forms, including performance on grades and conceptual measures by gender, differing attitudes and beliefs about physics (and the learning of physics) of male and female students, and documented gaps in the number of women recruited (and continuing) in the physics major. A primary research goal is to identify and quantify factors that strongly correlate with the gender gaps, and establish potential causal relationships. The research is informed by socio-cultural theory, and theories (including efficacy and stereotype threat) originating from social psychology. A major objective of the research program is to model the performance and trajectory of our students, to inform where interventions can be introduced with maximal impact for all students while decreasing the gender gap.

There are two driving research questions behind my work: 1) what factors can be documented to correlate with student performance in: conceptual learning (as measured by research-based assessments), attitudes and beliefs about science (as measured by research-based surveys), course performance and retention? 2) How do these factors differ for students retained in the physics major and those who enroll in introductory physics but do not continue? I hypothesize, based on my previous and ongoing work, that a number of factors, specifically math and physics background, incoming beliefs about the nature of physics, self-efficacy, affective factors and stereotype threat, can all be used to model the differential performance and retention of women in physics.

During the past year, while I was supported by the Chancellor’s Award, I made progress towards answering the driving research questions in the following ways: expansion and analysis of the data set to include Physics 2, analysis of survey data on self-efficacy, and a one-semester stereotype threat intervention experiment.

EXPANSION AND ANALYSIS OF THE DATA SET TO INCLUDE PHYSICS 2

Much of my previous work (Kost, et. al., 2009) and work from other researchers has focused on gender differences in the first-semester, calculus-based mechanics course. By studying gender differences in a different context, that of Physics 2, I am able to explore what contextual factors (like student population and student familiarity with the content) impact the gender disparities that I observe in these courses. I studied ten semesters of Physics 2 that were taught by seven different instructors. There are almost 4000 students included in this data set. All instructors except one were male.

Students were given the Brief Electricity and Magnetism Assessment (Ding, et. al., 2006) at the beginning and end of the semester (except in two semesters where the BEMA was only given at the end of the course) as a measure of their conceptual understanding of electricity and magnetism (E&M). As shown in the figure below, the gender differences on the pre-BEMA were small. On average, females scored about 1.5% lower than males on the BEMA pretest (effect size is 0.17). The gender gaps on the BEMA post-test were significantly different from zero in all semesters, and ranged from about 4% to about 10% (average effect size is 0.39). Despite males and females coming into Physics 2 with the same level of E&M conceptual understanding, at the end of the
semester, males are significantly outperforming females. It appears that a gender gap is created over the course of Physics 2. Another interpretation is that the pre-test gender gap is masked by students’ lack of familiarity with E&M content. When a conceptual survey of mechanics is used (from Physics 1) as a pre-course measure, we find that the gender gaps are reduced.

In addition to looking at conceptual survey performance, I also examined other course measures, like exam, homework, participation and final grades. I find that on average females outperform males on participation and homework, and males outperform females on exams. These differences offset one another such that the final course grades of males and females are not significantly different. There was one semester in which I did not observe this trend, in Spring 08 when there was a female instructor. In this semester, females outperformed males on homework and participation, and male and female exam scores were not significantly different. This resulted in females having course grades 1/3 of a letter grade higher than males. This was the only semester where we observed this trend. This suggests that having a female instructor can reduce gender disparities, but more work needs to be done to understand the impact of a female instructor before we can make that claim.

I was also able to track students from Physics 1 to Physics 2 to look at retention of physics majors. I found that more female physics majors than male dropped out of the physics sequence and changed their major to something other than physics between Physics 1 and Physics 2. These small differences resulted in a significant difference in the percent of male and female physics majors in Physics 2. More work needs to be done to follow students further along through the physics major and to gather more information about why female students are leaving the major.

Using multiple regression to model students’ post-BEMA scores, I found that prior physics understanding (both of mechanics and E&M), prior math performance (on standardized tests), and incoming attitudes and beliefs about physics all impact the gender gap. By controlling for these pre-measures, we can account for about 60% of the gender gap in post-BEMA scores. This model suggests that math and physics background and attitudes about physics are all important factors in understanding the gender gap in conceptual survey performance that we observe in Physics 2.
During the fall and spring of last year I wrote up these results for the Physical Review Special Topics: Physics Education Research journal; the paper is currently in review. I will be presenting these results at the American Association of Physics Teachers Summer Meeting (July 2010).

ANALYSIS OF SURVEY DATA ON SELF-EFFICACY

Self-efficacy (Bandura, 1977) characterizes the beliefs people have about their ability to complete the tasks required to be successful. Self-efficacy beliefs are thought to influence the choices that people make and the effort that they put forth towards being successful at some task. Several researchers have documented gender differences in self-efficacy (Pajares, 1996), and I hypothesized that self-efficacy may be a factor that influences male and female performance in physics. To that end, I administered a self-efficacy survey to students in both Physics 1 and Physics 2 at the beginning and end of the semester. I’m interested in three questions: 1) Are there gender differences in students’ physics self-efficacy at the beginning and/or end of the physics course? 2) Does students’ physics self-efficacy change over the course of the physics class? 3) Is physics self-efficacy a factor that can be used to model the differential performance of males and females?

Based on preliminary results, I’ve found that there are some gender differences in self-efficacy. While there are no gender differences in students’ overall physics self-efficacy, I do find significant gender differences on questions relating to students’ confidence regarding testing and performance. As is seen in the figure below, at the beginning of the semester, female students report being significantly less confident that they can complete tasks such as performing well on exams, demonstrating what they know on exams, defending their physics ideas to their peers, and learning physics concepts. I also find that both male and female students shift toward lower self-efficacy over the course of physics. Using regression modeling, I’ve begun to look at whether self-efficacy is a factor that affects student performance. Most interesting is whether self-efficacy can account for some of the gender differences that we observe, beyond those factors that we’ve already identified (math and physics background and incoming attitudes about physics). Preliminary modeling suggests that self-efficacy may be a useful predictor of student conceptual performance and may account for a fraction of the gender gap. These results will be presented at the American Association of Physics Teachers Summer Meeting (July 2010).
STEREOTYPE THREAT INTERVENTION EXPERIMENT

Stereotype threat is, “the threat of being viewed through the lens of a negative stereotype, or the fear of doing something that would inadvertently confirm that stereotype” (Steele, 1999). This fear of confirming the stereotype can inhibit performance and result in lower test scores for those students who are part of a stereotyped group, for instance, women in science. Researchers have demonstrated that stereotype threat can be alleviated through the use of self-affirmation (Cohen, et. al., 2006), where one affirms their self-worth and integrity. Based on prior work in the physics class, including survey questions that showed that female students were more stressed about exams than male students, I hypothesized that stereotype threat was at play in our courses. I decided to run an experiment to see if self-affirmation was effective in alleviating the gender disparities that I usually observe on course exams and standardized conceptual surveys.

During the first week of the physics course, students were randomly assigned to complete either a self-affirmation exercise (where students wrote about their own values) or a control exercise (where students wrote about someone else’s values). In the fourth week of the course, students again completed the same self-affirmation or control exercise. Scores on each of the four multiple-choice course exams were standardized and then averaged to get an average exam score for each student. I used a multiple regression analysis to test the effect of the self-affirmation exercise on students’ average exam scores. I was specifically interested in the interaction between condition and gender. The regression model also included a measure of prior mathematics ability as a covariate, and interactions with students’ responses to the statement: According to my own personal beliefs, I expect men to generally do better in physics than women.

I find that the gender x condition interaction is significant (p<0.01). As is seen in the figure below, the gender difference on exams in the control group is 0.62 standard deviation units (simple main effect is significant, p<0.001), while the gender gap in the self-affirmation group is 0.14 standard deviation units (simple main effect is not significant, p>0.1). Females in the self-affirmation group have average exam scores 0.31 standard deviation units above females in the control group (simple main effect is significant, p<0.02).

I also find that the three-way interaction between gender, condition, and belief in the gender stereotype is significant (p<0.01). The effect of self-affirmation is moderated by students’ belief in the stereotype that men do better in physics. As is seen in the figure below, males’ belief in the stereotype does not impact their average exam score.

![Graph showing effect size of gender difference on exams](image-url)
However, females who hold the belief that men should do better in physics and are in the control group have lower average exam scores than females in the self-affirmation group who endorse the stereotype. The self-affirmation exercise differentially benefits those females who believe that men are expected to do better in physics.

Similar trends are observed on the standardized conceptual survey as well, that is, the gender gap in the self-affirmation group is reduced compared to the control group. We have yet to examine other outcomes in the course, like self-efficacy. These results will be presented at the Physics Education Research Conference this summer (July 2010) and written up for the conference proceedings. Later this summer I will write up a full report to be published later this year.

REFERENCES