Gender Differences in Physics 1: The Impact of a Self-Affirmation Intervention

Lauren E. Kost-Smith*, Steven J. Pollock*, Noah D. Finkelstein*,
Geoffrey L. Cohen†, Tiffany A. Ito‡ and Akira Miyake‡

*Department of Physics, University of Colorado at Boulder, Boulder, Colorado 80309 USA
†School of Education and Department of Psychology, Stanford University, Stanford, California 94305 USA
‡Department of Psychology and Neuroscience, University of Colorado at Boulder, Boulder, Colorado 80309 USA

Abstract. Prior work at CU-Boulder has shown that a gender gap (difference in male and female performance) exists in both the pre- and post-course conceptual surveys, despite the use of interactive engagement techniques [Kost, et al., PRST-PER 5, 010101]. A potential explanation for this persistent gap is that stereotype threat, the fear of confirming a stereotype about one self, is inhibiting females' performance. Prior research has demonstrated that stereotype threat can be alleviated through the use of self-affirmation, a process of affirming one's overall self-worth and integrity [Cohen, et al., Science 313, 1307]. We report results of a randomized experiment testing the impact of a self-affirmation exercise on the gender gap in Physics 1. The gender gap on a conceptual post-survey is reduced from 19% for students who did not affirm their own values, to 9% for students who completed two 15-minute self-affirmation exercises at the beginning of the semester.

Keywords: gender, stereotype threat, self-affirmation, conceptual learning, introductory physics
PACS: 01.40.Fk, 01.40.G-, 01.40.gb, 01.75.+m

INTRODUCTION

Despite equal representations of males and females in biology, chemistry, and mathematics at the undergraduate level, females continue to make up only 21% of bachelor's degrees awarded in physics [1]. Our prior work has begun to address this disparity by examining gender differences in the first-semester, calculus-based mechanics course (Physics 1) [2]. Collecting data from twelve offerings of Physics 1, we find consistent pre- and post-course gender gaps (differences in male and female performance) on a conceptual survey of mechanics [3]. On average, the pre-course gender gap is about 10% and the post-course gender gap is about 12% (effect sizes of 0.4 to 0.5). These gender gaps exist despite the use of interactive engagement methods (e.g. Peer Instruction [4] and *Tutorials in Introductory Physics* [5]).

We have also found that background differences of males and females (differences in pre-course physics and math performance and incoming attitudes and beliefs) can account for about 70% of the post-course gender gap [2]. This suggests that the gender gap we observe at the end of Physics 1 is largely due to the under-preparation of females compared to males.

Currently, we are interested in understanding the mechanism by which the gender gap persists and what other factors impact the gender gap. Based on our prior work on self-efficacy, showing that females are more worried and nervous about taking exams [6], we hypothesize that stereotype threat may be inhibiting females' performance in the course. Researchers have demonstrated that stereotype threat can be alleviated through self-affirmation [7,8]. In this paper, we report on the results of a study to test the impact of a self-affirmation exercise on the gender gap in Physics 1. We find that the gender gap for students who affirmed their personal values was reduced by about half compared to students who did not affirm their own values. Further, when accounting for pre-course physics performance and students' endorsement of the stereotype, we find that the gender gap among students who affirmed their values was not significantly different from zero.

STEREOTYPE THREAT AND SELF-AFFIRMATION

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" [9]. This fear of confirming the stereotype can negatively impact members of a stereotyped group and result in worse performance. Spencer, et al. demonstrated the effect of stereotype threat by looking at females' performance on a difficult standardized math test [10]. They found that when students were told that the math test generally produced gender differences, females performed worse than equally qualified males. However, when students were told that the test generally did not show gender differences (alleviating females' concerns that their poor performance would confirm the stereotype that "girls can't do math"), females' performance increased, matching that of the males.

Several researchers have found that stereotype threat can be alleviated through self-affirmation [7,8]. Self-affirmation is a process through which a person affirms their overall sense of self-worth and integrity [11]. Work on self-affirmation assumes that people are motivated to maintain a positive sense of overall integrity, identity and worth. When our integrity or identity is threatened, we seek ways to resolve the threat. Because it is often difficult (or even impossible) to resolve the specific identity threat (that "girls can't do science"), individuals can affirm a more general sense of worth and integrity ("I'm a good person") or a specific, but unrelated, aspect of their identity ("I'm good at music"), which will help to protect them from the threat. In a laboratory experiment with college students, Martens, et al. [8] found that females who were given the opportunity to write about a characteristic of themselves that they valued performed better on a subsequent difficult math test than females who wrote about a characteristic that they did not value. Similar results were found by Cohen, et al. [7] comparing the school achievement of middle-school African-Americans who did and did not affirm their personal values.

STUDY DESIGN

We conducted a randomized experiment in order to test the impact of self-affirmation on the performance of males and females in Physics 1. Physics 1 is the first semester of the three-semester introductory physics sequence for science majors and engineers. It is a calculus-based course that covers Newton's laws, work, energy, momentum, and waves. In this semester, there were about 600 students in the class. Peer Instruction [4] and ConcepTests [4] were employed during lecture and students worked through *Tutorials in Introductory Physics* [5] during recitation. In terms of the curriculum, the course was nearly identical to previous semesters of Physics 1. The lead instructor of the course was experienced in interactive engagement methods and had recently taught this same course.

In the first week of the course, during recitation, students were randomly assigned to complete either a self-affirmation exercise where they wrote about values that were important to them, or a control exercise, where they wrote about values that were important to others. This writing exercise took about 15 minutes, and then students completed the Force and Motion Concept Evaluation (FMCE) [3] for the remainder of the 50-minute recitation. In the fourth week of the semester, the week before the first midterm exam, students again completed the same self-affirmation or control writing exercise as part of an online homework assignment. Students took three midterm exams and a final exam over the course of the semester. They completed the FMCE again during the last recitation. Exam and FMCE scores were collected as well as homework and participation scores.

In addition to the writing exercises, students were also asked to fill out an online, optional survey (worth token extra credit) in the second week of the course. The survey asked students several questions about their perception of the stereotype that men were better at physics than women. This survey was meant to measure students' awareness and endorsement of the gender stereotype.

We are interested in the impact of the self-affirmation intervention on student performance on the FMCE. We hypothesize that females who completed the self-affirmation exercise will perform better on the post-FMCE than females who completed the control exercise. Also, because we expect that there will be no significant impact of the self-affirmation exercise on male performance, we expect that the gender gap among students who completed the self-affirmation exercise to be smaller than the gender gap among students who completed the control exercise.

RESULTS

Students were only included in the analyses if they satisfied the following conditions: 1) they completed both writing exercises, 2) they took both the pre- and post-FMCE and the final exam, 3) they completed the stereotype threat survey, and 4) they had taken either the SAT- or ACT-Math test. This left us with a total of 308 students (52% of the class) [12]. The self-affirmation group had 137 males and 55 females, while the control group had 75 males and 41 females. By design, there are more students in the self-affirmation than the control condition (60% versus 40%). This was done to ensure that more students would receive the potentially beneficial self-affirmation exercise. There were no significant

differences on any prior factors [13] between the self-affirmation and control groups, by gender.

We used a multiple regression analysis to test the effect of the self-affirmation exercise on students' post-FMCE scores. This standard statistical approach allowed us to create a model of FMCE post-scores that accounts for other factors in addition to gender and experimental condition. In these models, we were specifically interested in the interaction between gender and condition. The significance of this interaction indicates that the gender gap in the self-affirmation condition is not equal to the gender gap in the control condition. We tested three consecutive models. The first included only gender, condition, and the gender×condition interaction term. In the second model, we included FMCE pre-score as a covariate [14]. In the final model, we test the three-way interaction between gender, condition, and students' belief in the stereotype. All three models are shown in Table 1. In each model, gender is coded as 1 for females and 0 for males, and condition is coded as 1 for affirmation and 0 for control. FMCE pre-score and stereotype belief are both centered (mean=0).

TABLE 1. Coefficient estimates and multiple regression model statistics for each regression model. The * (**) indicates the coefficient is significant at the 0.05 (0.01) level.

	Model 1	**Model 2**	**Model 3**
Model-level statistics			
Multiple R^2	0.053	0.328	0.353
F statistic p value	0.001	<0.001	<0.001
Res. Std. Err.	27.22	22.96	22.68
Predictors	**Coeff.**	**Coeff.**	**Coeff.**
Intercept	74.7**	71.1**	71.6**
Gender	-18.6**	-8.3	-12.7**
Condition	-2.1	-0.7	-0.6
Gender × Condition	9.6	6.9	12.9*
FMCE Pretest		0.6**	0.6**
Stereotype Belief (ST)			-2.7
Gender × ST			-7.2
Condition × ST			0.7
Gender × Condition × ST			13.3*

In all three models, the variables included account for a significant fraction of the variance in FMCE post-scores (F statistic p value < 0.01). Looking first at Model 1, we see that the gender×condition interaction is not significant (p=0.16), meaning that the gender gaps in the affirmation and control groups are not significantly different. Though the difference is not statistically significant, the gender gap in the control group is 19% ± 5%, while the gender gap in the self-affirmation group is 9% ± 4%. The gender gap in the self-affirmation group is about one-half that of the control group.

In Model 2 we add the FMCE pretest as a covariate to control for students' pre-course physics performance. This allows us to compare prototypical students who have the same FMCE pretest score. We find again, that the gender×condition interaction is not significant (p=0.24). Looking at the gender gaps in each condition, we find that for students with the average FMCE pretest score (35.3%) the gender gap in the control group is 8% ± 5%, while the gender gap in the self-affirmation group is 1% ± 4%. Again, we see a large (but not statistically significant) reduction of the gender gap in the self-affirmation condition compared to the control condition.

Model 3, the final model, includes a measure of students' responses to the following statement: *According to my own personal beliefs, I expect men to generally do better in physics than women*. Students agreed or disagreed with the statement on a 5-point Likert scale. Figure 1 shows the distribution of male and female responses to the statement.

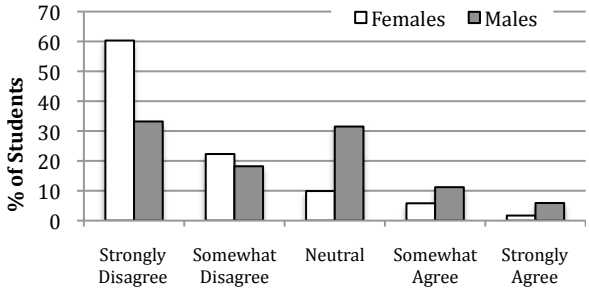

FIGURE 1. Distribution of student responses to the statement: *According to my own personal beliefs, I expect men to generally do better in physics than women*.

We tested the three-way interaction of gender×condition×stereotype belief. The significance of this three-way interaction indicates that the gender×condition interaction varies depending on how much students believe the stereotype. This is most easily seen in Figure 2. There is no relationship between belief in the stereotype and FMCE post-score for males in either the affirmation or control group (simple slopes not significantly different from zero [15]). However, for females, believing in the gender stereotype negatively impacts their FMCE post-score (simple slope is significant), unless they completed the self-affirmation exercises (slope not significant). This demonstrates not only that those who moderately endorse the stereotype and are part of the stereotyped group are harmed the most by stereotype threat, but also that the self-affirmation exercises were particularly beneficial for those students under the highest threat.

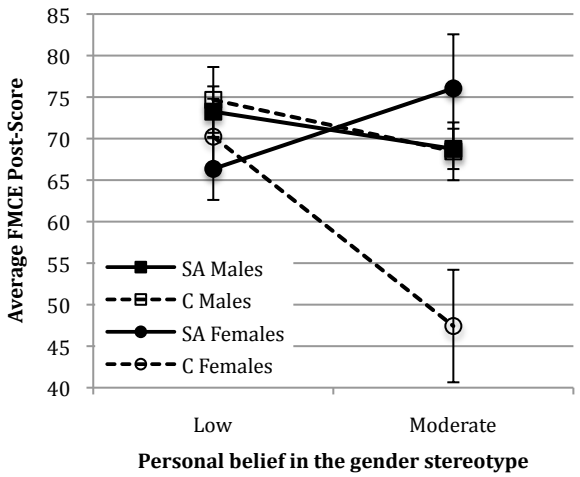

FIGURE 2. Predicted FMCE post-scores for levels of agreement with the gender stereotype (pre-FMCE=mean). SA indicates self-affirmation and C indicates control.

Using Model 3 we can estimate the gender gaps for the self-affirmation and control groups at the average FMCE pre-score and average level of belief in the stereotype. The gender gap in the control group is 13% ± 5% ($p<0.01$), while the gender gap in the self-affirmation group is 0% ± 4% ($p>0.9$). The gender gap has been eliminated in the self-affirmation condition. Recall that we expected that females in the self-affirmation condition would have higher FMCE scores than females in the control condition, and males' scores would not be different in the two conditions. We find that this is the case. Females in the self-affirmation condition have FMCE post-scores 12% ± 5% ($p<0.01$) higher than females in the control condition, and males in the self-affirmation condition have FMCE post-scores 1% ± 3% ($p>0.8$) lower than males in the control condition.

DISCUSSION

We have demonstrated that two simple, 15-minute writing exercises completed at the beginning of the semester can increase females' performance (while not significantly hurting male performance) on the FMCE post-survey and can reduce the gender gap. Further, the effect of the self-affirmation is moderated by students' belief in the stereotype that men will do better in physics than women. The self-affirmation is more beneficial for females who moderately endorse the stereotype, those females who are most threatened. We have shown that the self-affirmation exercises are effective in an authentic physics classroom environment for college-age students. These results are also confirmed in exam and course grades [16].

Our findings here are consistent with our prior work; a large fraction of the post-course gender gap can be accounted for by pre-course gender differences. Additionally, the reduction of the gender gap among affirmed students supports our hypothesis that the remainder of the gender gap is due, at least in part, to stereotype threat.

There are at least two implications to take away from this study: 1) We, as educators and researchers, need to be more aware of and attentive to psychological factors that can impact student performance in our courses, and 2) We need to do more to help those students who are under-prepared to succeed in introductory physics courses, a group that is predominantly female.

ACKNOWLEDGMENTS

This work is supported by iSTEM (#0833364) and DRL (#0910373). We thank the PER at Colorado group and the faculty that were involved. We also thank Lindsey Newnes and Natalie Golaszewski for their assistance with data collection.

REFERENCES

1. NSF, Division of Science Resources Statistics, NSF 09-305, Arlington, VA, 2009. www.nsf.gov/statistics/wmpd/
2. L. E. Kost, et al., *Phys. Rev. ST Phys. Ed. Res.* **5**, 010101 (2009).
3. R. K. Thornton and D. R. Sokoloff, *Am. J. Phys*. **74**, 118 (2006).
4. E. Mazur, *Peer Instruction: A User's Manual*, Upper Saddle River, NJ: Prentice-Hall, 1997.
5. L. C. McDermott and P. S. Schaffer, *Tutorials in Introductory Physics*, Upper Saddle River, NJ: Prentice-Hall, 2002.
6. L. E. Kost, et al. in *2009 PERC Proceedings*, ed. by M. Sabella, et al., AIP Conference Proceedings 1179, AIP, Melville, NY, 2009, pp. 177.
7. G. L. Cohen, et al., *Science*, **313**, 1307 (2006).
8. A. Martens, *J. Exp. Soc. Psych*. **42**, 236 (2006).
9. C. M. Steele, *Am. Psych*. **52**(6), 613 (1997).
10. S. J. Spencer, et al., *J. Exp. Soc. Psych*. **35**, 4 (1999).
11. C. M. Steele, in *Advances in Experimental Social Psychology*, ed. by L. Berkowitz, New York: Academic Press, 1988, pp. 261.
12. As is usually the case, students who participated in the study had higher course grades than those who did not.
13. Prior factors included SAT-Verbal, SAT-Math, SAT total score, ACT-Reading, ACT-English, ACT-Math, ACT-Scientific Reasoning, ACT total score, years of high school physics and calculus, and high school GPA.
14. The gender×condition interaction was not significant for the FMCE pre-test ($p=0.47$), indicating that there was no effect of the intervention on the FMCE pre-test.
15. L. S. Aiken and S. G. West, *Multiple Regression: Testing and Interpreting Interactions*, Thousand Oaks, CA: Sage, 1991.
16. Results in preparation.