STEREOTYPE THREAT IN THE INTRODUCTORY SCIENCE CLASSROOM: INVESTIGATING ITS EXISTENCE AND TRIGGERS

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Shanda Deleen Lauer

In Partial Fulfillment
for the Degree of
MASTER OF SCIENCE

Major Department:
Biological Sciences

June 2013

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Stereotype Threat in the Introductory Science Classroom: Investigating its Existence and Triggers

By

Shanda Deleen Lauer

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Jennifer Momsen
Chair

Dr. Erika Offerdahl

Dr. Erin Gillam

Dr. Lisa Montplaisir

Approved:

June 10, 2013 — Date

Dr. Wendy Reed — Department Chair

ABSTRACT

Research in undergraduate education has documented achievement gaps between men and women in math and physics that may reflect, in part, a response to perceived stereotype threat. My research efforts aimed to reduce achievement gaps by mediating the impact of stereotype threat in introductory science classrooms with a short, values-affirmation writing exercise. The purpose of this research was to (1) investigate and compare the performance of women and men across introductory science sequences (biology, biochemistry, physics), (2) document endorsement of stereotype threat, (3) investigate the utility of a values-affirmation writing task in reducing achievement gaps, (4) provide a meta-analysis of triggers causing stereotype threat, and (5) advise classroom practices to avoid stereotype threat. In this study, analysis of final grades and normalized learning gains on concept inventories revealed no achievement gap in the courses sampled, little stereotype threat endorsement, and no impact of the values-affirmation writing task on student performance.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under NSF-HRD 0811239 as part of the NDSU Advance/FORWARD program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. STEREOTYPED: INVESTIGATING GENDER IN INTRODUCTORY SCIENCE COURSES[1]

Introduction

Despite decades of active recruitment, women remain underrepresented in science, technology, engineering, and math (STEM) disciplines both in the United States and globally (Hewlett *et al.*, 2008; Simard *et al.*, 2008). Women leave STEM fields at all stages of their careers – as undergraduates, graduate students, professionals, and in the transitions between each stage, a phenomenon described as the leaky pipeline. In biology, for example, although women have reached parity with men when graduating from undergraduate and post-graduate schooling, women comprise approximately 1/3 of the academic workforce (National Science Foundation, Division of Science Resources Statistics, 2011). In contrast, the physics pipeline leak begins much earlier and is more substantial. Despite the fact that women and men are nearly equally represented in high school physics classes (44% vs. 56%) the pipeline turns into a "gaping hole" when they reach college (McCullough, 2002). Women comprise only 21% of physics undergraduate degrees, 22% of Masters and 16% of PhDs (Mulvey and Nicholson, 2008). As these women move into academe and professional roles, they comprise 11% of the workforce (National Science Foundation, Division of Science Resources Statistics, 2011).

The underlying causes of this disparity between men and women are numerous, complex, and pervasive. However, a recent meta-analysis of research on the gender gap in STEM (Hill *et al.*, 2010) found bias, stereotype threat, and social factors as prime driving forces contributing to the loss of women from STEM fields. In fact, recent work by Moss-Racusin *et al.* (2012), found

---

[1]The material in this chapter was co-authored by Shanda Lauer and Jennifer Momsen, Erika

science faculty across disciplines and regardless of gender exhibited an unconscious gender bias against undergraduate women, underscoring the pervasive and persistent nature of cultural stereotypes regarding women in science.

Gender and achievement in undergraduate science courses

The disparity between women and men in STEM disciplines may extend to achievement at the college level, resulting in a gender achievement gap - the persistent and pervasive underperformance of women as measured by exam scores, course grades, and learning gains on validated concept inventories.

Evidence for an achievement gap in biology and biochemistry at the undergraduate level is largely missing, in part because the fields are young. Women routinely underperform their male counterparts on the MCAT, a pattern that can be traced back at least a decade (American Association of Medical Colleges, 2012). Further, a recent study by Willoughby and Metz (2009) found mixed evidence of a gender gap in an introductory biology course: women had significantly lower normalized learning gains as measured by a biological diagnostic test, but this result was not reproducible with any other measure, including alternative learning gain calculations, overall course grades, and individual exam scores. Many students from introductory biology go on to take introductory biochemistry. Yet there are few diagnostic tests for biochemistry (e.g., American Chemical Society Biochemistry Exam, Biochemistry and Cell Biology GRE), and, to date, none have been used to explore the existence of a gender gap. Such limited results underscore the need for additional studies of how women and men perform in undergraduate life sciences courses, a need echoed by the recently released report on the status of discipline-based education research (DBER) by the National Academies of Science (2012).

In contrast, gender achievement gaps are well documented in physics at the undergraduate level (Lorenzo *et al.*, 2006; Pollock *et al.*, 2007; Kost *et al.*, 2009; Brewe *et al.*, 2010; Kost-Smith *et al.*, 2010). The calculus-based introductory physics sequence, a gateway to majors in physics and many other STEM disciplines, is the most frequently studied in Physics Education Research (PER). A distinct gender gap exists on conceptual surveys among students before instruction (Lorenzo *et al.*, 2006; Pollock *et al.*, 2007; Brewe *et al.*, 2010), but some of this disparity may be due to gender bias in the instruments themselves (McCullough and Meltzer, 2001; Docktor *et al.*, 2008; Willoughby and Metz, 2009; Dietz *et al.*, 2012). In courses with traditional instructional methods this gap appears to persist; however, when instruction consists of highly interactive, research-validated instruction, the prevalence of an achievement gap is less consistent. Although learning gains are significant regardless of gender, some research finds the achievement gap reduced (Lorenzo *et al.*, 2006), while other research finds the gap persists (Pollock *et al.*, 2007; Brewe *et al.*, 2010). As noted previously, the presence of an achievement gap may be an artifact of over-reliance on potentially biased conceptual surveys, especially when associated course grades and final exams do not reveal such a significant gap (Docktor *et al.*, 2008; Willoughby and Metz, 2009).

In many instances, the gender gap in physics is attributed to disparities in mathematical preparation and ability. While a strong and persistent belief in a gender achievement gap in mathematics has prevailed for decades (e.g., Kane and Mertz, 2012), evidence for its existence is less conclusive (e.g., Hyde, 2005; Guiso *et al.*, 2008). In a meta-analysis of six large survey studies, Hedges and Nowell (1995) documented a small mean difference in mathematics achievement between men and women and modest differences in variance. More recent data in the United States refute a mathematics gender achievement gap, at least in the general populace grades 2 through 11 (Hyde *et al.*, 2008). Analyses of international data collected through studies such as the 2003 Trends in International Mathematics and Science Study (TIMMS) and 2003

Programme for International Student Assessment (PISA) reveal significant variability between nations in the presence and effect size of a gap (Guiso *et al.*, 2008; Nosek *et al.*, 2009). While there seems to be some agreement that, in some contexts, the gender achievement gap is narrowing or may no longer exist, the implications for such a gap, no matter how small, are still of import. Hedges and Friedman (1993) predict that even a difference as small as 0.3 standard deviation coupled with modest variance can account for as much as 2.5 times as many men in the top scoring percentiles than women.

In instances where an achievement gap has been documented, the underlying causes of these differences in math performance are likely numerous and the relationships between them complex. Contextual factors play a key role in predicting differences in achievement. Analyses of TIMMS and PISA data identified socio-cultural indicators of gender equality within a nation as a strong predictor of differences in achievement (Guiso *et al.*, 2008; Nosek *et al.*, 2009). Neiderle and Vesterlund (2011) provide evidence that women perform differently than men on mathematics-related tasks when the situation is perceived to be highly competitive.

Stereotype threat

Stereotype threat, described as a “risk of confirming […] a negative stereotype about one’s group” (Steele and Aronson, 1995), may undermine achievement in the STEM classroom. Although not limited to gender – stereotype threat can apply to many intrinsic characteristics, including race, ethnicity, income level, and academic ability (Allport, 1954; Steele, 1997) – we focus here on the impact of stereotype threat on the performance of women in undergraduate STEM courses.

Stereotype threat may be highly contextual, triggered by a survey item (Steele and Aronson, 1995), the gender of the instructor (Delisle *et al.*, 2009), or instructional practices (Kreutzer and Boudreaux, 2012) and can undermine academic success in several ways. First, stereotype threat can produce stress and induce anxiety, causing a student to become more self-conscious about their performance and to actively try to suppress those emotions, which may tax working memory and lead to decreased performance (Steele and Aronson, 1995; Schmader *et al.*, 2008; Delisle *et al.*, 2009). Second, prolonged exposure to stereotype threat can result in disidentification, wherein a student stops associating with a given stereotyped group and avoids situations likely to be perceived as threatening (Aronson *et al.*, 2002; Steele *et al.*, 2002). In science, stereotype threat may contribute to the leaky pipeline, causing women to attrition from science-related majors.

While stereotype threat has become a popular explanation for differences in performance between men and women in STEM disciplines, recent work by Stoet and Geary (2012) calls to question the strength of empirical evidence supporting this hypothesis. They reviewed the research on gender differences in mathematics and performance and achievement to determine the strength of evidence supporting results from the original, critical study documenting activation of stereotype threat in mathematics (Spencer *et al.*, 1999). Stoet and Geary (2012) conclude that the evidence for activation of stereotype threat as the mediating factor of a gender achievement gap is far from robust. Although they identified 141 articles related to stereotype threat in mathematics, 20 of these were replication studies. Of these, just 11 (55%) were able to replicate the activation of stereotype threat as in the original paper. While they do not dismiss stereotype threat as a valid hypothesis, they do call to question the strength of the effect on achievement and performance and caution researchers and policy makers alike to consider the vast array of other possible contributing factors to the gender achievement gap.

Reducing the impact of stereotype threat

Empirical work focused on ways to reduce or eliminate the effects of stereotype threat has revealed a number of simple yet effective measures including educating at-risk populations

(Johns *et al.*, 2005) and manipulating test-taking instructions (Steele and Aronson, 1995; Spencer *et al.*, 1999; Johns *et al.*, 2005). Social psychologists have also reduced the effects through mediation of contextual and societal factors related to stereotypes. Individuation has proved effective by explicitly distinguishing between the stereotyped individual and the stereotype to minimize stereotype usage (Locksley *et al.*, 1980; Langer *et al.*, 1985) and allows stereotyped students to distance themselves from the stereotype in question while remaining engaged in the task or course (Ambady *et al.*, 2004). Finally, since women are more likely to endorse the stereotype that science is for men when suitable female role models are largely absent (i.e., few female faculty) (Delisle *et al.*, 2009), simply increasing the visibility of and engagement with positive female role models has proven efficacious (McIntyre *et al.*, 2004). In fact, simply having a competent woman administer a mathematics exam was sufficient to reduce the achievement gap in one study (Marx and Roman, 2002).

Values-affirmation tasks have recently received a great deal of attention (e.g., Cohen *et al.*, 2006; Miyake *et al*., 2010) for their ability to reduce or eliminate stereotype threat. In this type of intervention, individuals take 10 – 15 minutes to write about values that are personally important but unrelated to the course. Such writing tasks appear effective in reducing or eliminating stereotype threat for African-Americans (Cohen *et al.*, 2006; Walton and Cohen, 2007) and women (Martens *et al.*, 2006; Miyake *et al.*, 2010), with effects that may persist over time (Cohen *et al.*, 2009; Walton and Cohen, 2011). Although short and simple, values-affirmation writing tasks draw directly on students' experiences to actively engage each student as an individual (Yeager and Walton, 2011) and may promote deep processing to effect powerful results (Schwartz and Martin, 2004; Chase *et al.*, 2009). Thus, although simple, values affirmation writing tasks have the potential to profoundly impact students experiencing stereotype threat (Yeager and Walton, 2011).

Testing the efficacy of values-affirmation tasks in introductory science

The work of Miyake *et al.* (2010) and Cohen *et al.* (2006) is encouraging, but each study represents only a single course or cohort of students at one institution. Given the complex nature of the classroom and the myriad of factors that contribute to learning, it is necessary to replicate the values-affirmation study across institutions, semesters, and courses; indeed, this lack of replication studies is a serious deficit of current DBER practices (Singer *et al.*, 2012).

This study addresses this deficiency and specifically investigates the gender achievement gap across introductory science courses and tests the efficacy of a values-affirmation task in improving student performance. Specifically, we (1) characterized and compared the performance of women and men across three introductory science sequences (biology, biochemistry, and physics) at a large, public, research-intensive university, (2) documented endorsement of stereotype threat in these science courses, and (3) determined the utility of a values-affirmation writing task in reducing achievement gaps that may exist.

Methods

University and course context

This land-grant, research university serves over 14,000 undergraduate and graduate students. Women comprise 42% of the undergraduate population and 50% of the graduate population. Across the university, incoming freshmen have an average composite ACT score of 23.8, and an average high school GPA of 3.37.

This study targeted four science courses, considered introductory for majors in the discipline, including introductory calculus-based physics 1 and 2, introductory biology, and introductory biochemistry. Introductory physics 1 is a lecture-based course, taught by a male faculty member, and introduces Newtonian mechanics of translational and rotational motion, energy, work, power, momentum, conservation of energy and momentum, periodic motion, waves, sound, heat and thermodynamics. Enrollment is typically 90 – 100 students. Introductory physics 2, taught by a female faculty member, is also a lecture-based course, and focuses on conceptual understanding of topics including electric charge, electric field, potential and current, magnetic field, capacitance, resistance, and inductance, circuits, electromagnetic waves, and optics. Enrollment is typically around 200 students. Introductory biology is a very large (300 – 400 students) lecture-based course, taught by a female faculty member, and introduces students to cellular and molecular biology, genetics, and evolution. Biochemistry is also a large lecture-based course with average enrollments of 300 students, taught by a female faculty member, and focuses on biomolecules, generation and use of metabolic energy, biosynthesis, metabolic regulation, storage, transmission, and expression of genetic information.

Gender achievement gap

To investigate the presence and persistence of a gender achievement gap, we collected data, specifically final course grades by gender, from iterations of these courses taught in the 2010-11 academic year. We also collected these data from Fall 2011, the same semester in which the values-affirmation writing task was implemented.

Values-affirmation exercise

We followed the protocol described by Miyake *et al.* (2010) to implement the values-affirmation exercise in four different introductory science courses in the Fall 2011 semester. This exercise was unrelated to the content of any of the courses included in this study. The exercise was distributed in a double-blind fashion within the lecture component of each course. Given the predicted benefits of the task, we randomly assigned approximately 60% of each course to the values-affirmation treatment group and 40% to the control group (Table 1). The first writing exercise was distributed the second week of classes, following students' completion of a discipline-appropriate concept inventory (Figure 1). A research assistant unaffiliated with any of the courses included in this study implemented the writing task following a well-defined script. Students were given 15 minutes to complete the writing task.

Table 1. **Participants in the values-affirmation task, as distributed among treatment groups.**

	Males (T/C)	**Females (T/C)**	**Total**
Introductory biology	138 (74/64)	131 (85/46)	269
Biochemistry	97 (61/36)	122 (74/48)	218
Physics 1	52 (29/23)	13 (9/4)	65
Physics 2	111 (66/45)	15 (9/6)	126

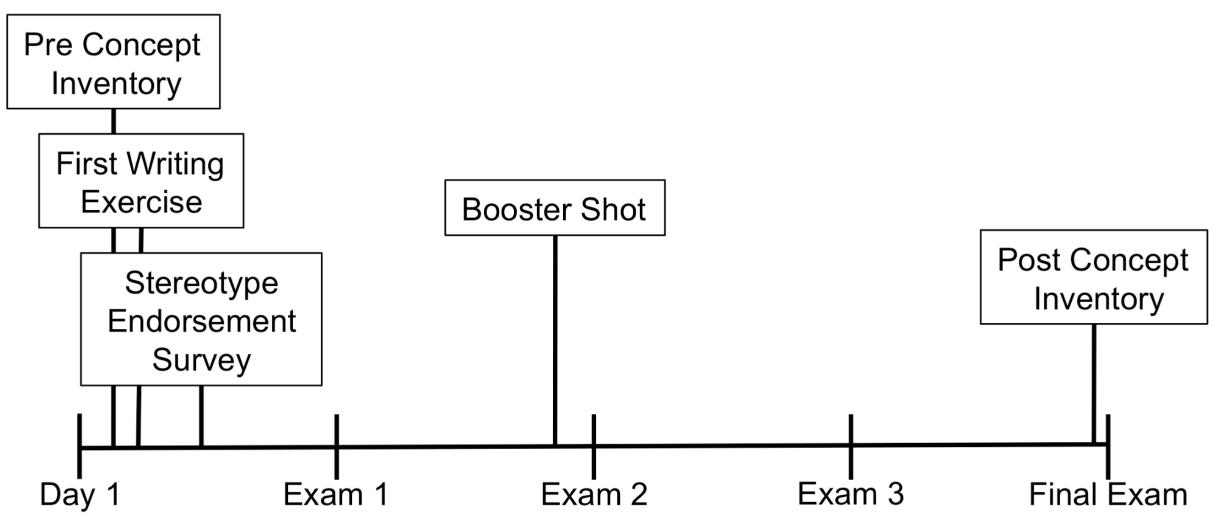

Figure 1. General timeline of the intervention and data collection.

In the week prior to the second exam, students were asked to again complete the values-affirmation writing exercise. This 'booster-shot' was intended to help students reaffirm their values. This time, the writing exercise was administered online through a class webpage as a regular homework assignment. Students were invited individually to follow a link to an online replica of the writing exercise done in class, and the treatment conditions were kept the same as the first implementation. The instructions were the same, suggesting that students spend approximately 15 minutes on the exercise.

Stereotype endorsement measures

Again, following the protocol of Miyake *et al.* (2010), we also distributed a survey to measure students' endorsement of gendered stereotype threats, namely that men are generally better at a particular science (e.g., physics, biochemistry, or biology). Within the 45 item survey, we distributed two stereotype endorsement questions, customized to each course: (1) according to my own personal beliefs, I expect men to generally do better than women in physics (or biochemistry or biology), and (2) according to my own personal beliefs, I expect women to generally do better than men in physics (or biology or biochemistry). The participants were asked to indicate their agreement on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). This approach does not specifically prime students' stereotype threat (e.g., by asking them to identify as female); rather, stereotype threat is activated by situational pressure, that is, being aware of the stereotype threat and being a member of the threatened group (e.g., women perform more poorly than men in science and I am a woman) (e.g., Marx and Stapel, 2006).

Outcome measures

The main outcome measures for this study included final course grades and learning gains (Hake, 1998), the latter measured by student performance on a discipline-appropriate concept inventory (Table 2). To test for differences between the performance of men and women, we use a Chi-squared analysis, with Fisher's exact test when sample sizes were too small to meet the assumptions of the Chi-squared analysis. To compare learning gains of men and women in treatment and control groups, we used Student's t-test. Where appropriate, we calculated effect sizes using Cohen's *V* or *d* and include confidence intervals. Analyses were conducted using SAS software.

Table 2. **Discipline-specific concept inventories.**

Course	**Concept Inventory**
Physics 1	Force and Motion Conceptual Evaluation[1]
Physics 2	Brief Electricity and Magnetism Assessment[2]
Introductory biology	Concept Inventory of Natural Selection[3]
Introductory biochemistry	Introductory Molecular and Cell Biology Assessment[4]

[1]Thornton, 1998.[2]Ding *et al.*, 2006. [3]Anderson *et al.*, 2002. [4]Shi *et al.*, 2010.

Results

Gender achievement gap

There was no significant relationship between the distribution of final course grades and gender in biology or physics for any semester or section (Table 3). For biochemistry, however, there was significance, which shows a relationship between gender and letter grade for Fall 2011; however, women seemed to outperform men in this class and semester, although the effect size was small ($V = 0.2$, 95% CI [0.14, 0.3]). Further, we found no significant differences between normalized learning gains of men and women for any course (Table 4).

Stereotype threat endorsement

In all courses, students overwhelmingly rejected the claim that men do better than women in biology, biochemistry, or physics, with more than 2/3 of students strongly disagreeing or disagreeing with the statement (Figure 2). The distribution of responses for men differed significantly from women only in biology ($\chi^2(4) = 23.29$, $p < 0.001$), with women more likely to disagree with this claim.

Values-affirmation writing task

In all courses but one, physics 2, learning gains were higher for the treatment group over the control group, significantly so for only physics 1 (Table 5), with a moderate effect size (d = -0.7, 95% CI [-1.3, -0.09]). Further, in all courses but physics 1, final course grades were higher for the control group over the treatment group, significantly so for only physics 2 (Table 6), although the effect size was small (d = 0.4, 95% CI [0.04, 0.8]). Further, there was no significant difference in the distribution of final grades between treatment and control groups for women or men in any course (Table 7).

Table 3**. Chi-squared analysis of final course grade distributions by gender.**

Course	**Year**	**df**	**n**	χ^2	**p-value**
Introductory biology	2010	4	323	1.83	0.78
	2011	4	269	5.06	0.28
Biochemistry	2010	4	264	2.26	0.69
	2011	4	219	10.05	0.04
Physics 1	2010	4	74	3.14	0.56[1]
	2011	4	65	5.41	0.27[1]
Physics 2	2010	4	188	2.52	0.71[1]
	2011	4	126	1.28	0.94[1]

[1]Fisher's exact test used when the data set violated the assumption that each expected cell count was greater than five.

Table 4. **Comparing normalized learning gains for men and women in fall 2011.**

Course	**Mean difference[1]**	**Df**	**t**	**p-value**
Introductory biology	-0.01	171.18	-0.20	0.84
Biochemistry	0.01	183	0.14	0.89
Physics 1	-0.17	42	-1.27	0.21
Physics 2	-0.10	89	-1.46	0.15

[1]A negative mean difference value indicates higher learning gains for the treatment group.

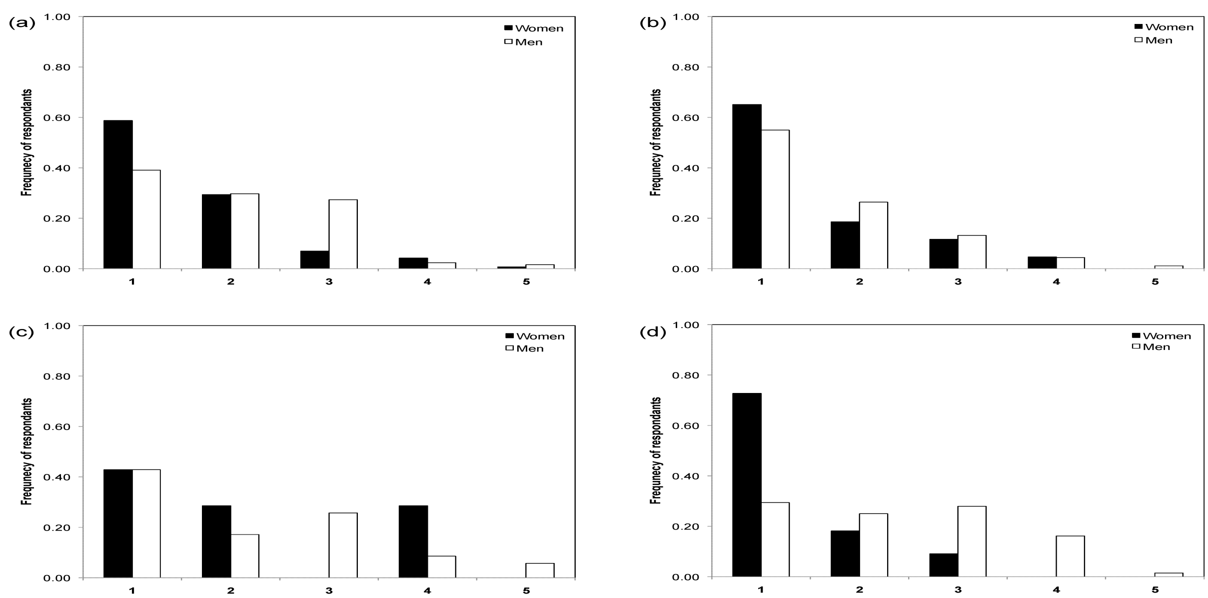

Figure 2. Frequency of student responses to the question, I expect men to generally do better than women in (a) biology (n=227), (b) biochemistry (n=243), (c) physics 1 (n=44), or (d) physics 2 (n=91), where 1 = strongly disagree to 5 = strongly agree.

Table 5. **Comparison of normalized learning gains between treatment and control groups.**

Course	**Mean difference[1]**	**df**	**t**	**p-value**
Introductory biology	-0.07	130.13	-0.90	0.37
Biochemistry	-0.06	183	-1.36	0.18
Physics 1	-0.25	42	-2.32	0.03
Physics 2	0.04	87.36	0.83	0.41

[1]A negative mean difference value indicates higher learning gains for the treatment group.

Table 6. **Comparison of final course grades between treatment and control groups.**

Course	Mean difference	df	t	p-value
Introductory biology	1.70	257.94	0.96	0.34
Biochemistry	1.50	209.84	1.07	0.29
Physics 1	-3.96	63	-0.82	0.42
Physics 2	5.44	121.76	2.22	0.03

Table 7. **Comparison of final course grades between treatment and control groups.**

Course	Gender	Mean (± St Dev)	df	n	χ^2	p-value[1]
Introductory biology	F	74.3 ± 15.1	4	131	7.67	0.11
	M	73.7 ± 14.4	4	138	3.03	0.57
Biochemistry	F	80.0 ± 8.8	4	122	2.21	0.80
	M	77.4 ± 12.7	4	97	6.55	0.17
Physics 1	F	83.2 ± 10.8	4	13	5.33	0.13
	M	78.8 ± 20.7	4	52	1.72	0.80
Physics 2	F	82.9 ± 11.1	4	15	2.85	0.60
	M	80.2 ± 15.5	4	111	3.28	0.55

[1]Fisher's exact test used when the data set violated the assumption that each expected cell count was greater than five.

Discussion

The existence of an achievement gap is often an assumption of the undergraduate physics classroom, yet remains an unknown in introductory biology and biochemistry courses. However, across semesters and outcome measures, we found no substantial evidence of an achievement gap between men and women in either introductory calculus-based physics courses or introductory biology and biochemistry. Although these findings align with studies in astronomy

(Hufnagel *et al.*, 2004; Willoughby and Metz, 2009) and biology (Willoughby and Metz, 2009), they contradict what is typically reported in physics (Lorenzo *et al.*, 2006; Pollock *et al.*, 2007; Miyake *et al.*, 2010). Such discrepancies may be attributable to biases in how learning gains are calculated; indeed, normalized learning gains are particularly susceptible to bias since there is a strong relationship between pretest scores and normalized learning gains (Coletta and Phillips, 2005; Brogt *et al.*, 2007). For example, since men typically have higher pre-test scores than women on common physics concept inventories (e.g., FCI or FMCE), the subsequent calculation of normalized learning gains is particularly likely to identify a gender achievement gap. Our results utilized normalized learning gains, which further underscores the lack of an achievement gap in the sampled science courses.

Explaining gender achievement gaps, however, goes beyond statistical biases. Stereotype threat can play a role in student achievement, especially, as noted, on standardized tests and concept inventories in science and math. Women in science often ascribe to a negative stereotype regarding women's scientific competency. However, in this study we found little to support the claim that women in the sampled population were endorsing a stereotype threat; rather, our evidence suggests that most women, and even men, reject this claim. We are cautious in our interpretation of these data for several reasons. In physics, these results may reflect the small sample size of women, although in such cases we might expect women would more readily self-identify as female and thus face an increased risk of experiencing stereotype threat. However, these results may reflect a stereotype reactance effect, wherein the stereotype is so blatant that women respond by over performing (Kray *et al.*, 2001). Although our sample sizes for introductory biology and biochemistry are more robust, we believe this study is one of the first to explicitly explore gender achievement gaps and stereotype threat at the undergraduate level in either biology or biochemistry. As such, this research represents a single time point and institution and is hardly representative of national trends.

Still, these results are perplexing in light of the broader research landscape, prompting us to question why these students may not ascribe to gender-based stereotype threats. One possible explanation emerges from self-efficacy literature, specifically, the role of vicarious experiences in shaping student's beliefs regarding self-efficacy. Vicarious experiences involve more than just a positive role model; they reflect repeated observations of "others perform[ing] threatening activities without adverse consequences" (Bandura, 1977). By extension, the observer can predict that her hard work and persistence can result in success. In the undergraduate setting, vicarious experiences for women include observing women in roles of authority and as experts, such as lab and recitation TAs and course instructors. Given the institutional context of this study, vicarious experiences may play an important role in student's perception of self-efficacy and stereotype threat. Introductory biology and biochemistry are both taught by female instructors and female graduate students often lead the associated labs; thus, students are afforded multiple opportunities to observe women doing biology and biochemistry and may have greater self-efficacy when doing biology and biochemistry themselves. All women enrolled in biochemistry would have successfully completed at least one course in biology, and many would have also successfully completed a physics course. Prior success in biology and physics might serve to affirm women's beliefs in biochemistry that they "belong" in the field. Conversely, the physics department has only one female faculty member, and at the time of this study, no female graduate students. Thus, opportunities to observe women performing "threatening activities" were rare. However, we note the somewhat anomalous result of physics 2, where 91% of women disagree or strongly disagree with the claim that men generally do better in physics. Taught by a female faculty member, instruction in this course regularly offers women an opportunity to observe a woman *doing* physics and may promote positive feelings of self-efficacy in female students. Further, women enrolled in physics 2 had successfully completed physics 1 (or equivalent), which is a prerequisite to physics 2, and therefore may have already identified themselves as capable of doing well in physics.

Just as vicarious experiences can influence endorsement of stereotype threat, other contextual elements might explain our inability to detect meaningful differences in achievement and stereotype threat endorsement. Schmader *et al.* (2008) presented a model postulating a link between stereotype threat and the activation of processes that tax otherwise available cognitive resources (e.g., physiological stress, suppression of negative emotions, and performance monitoring). When an individual endorses a stereotype, they are less likely to perform well because they have fewer cognitive resources available. Alter *et al.* (2010) demonstrate that the way in which a task is presented can affect the degree to which an individual endorses or identifies with a given stereotype. They demonstrated differential performance in stereotyped groups dependent upon how a task was presented – either as a task or as a challenge. When groups susceptible to stereotype threat were presented a task couched as a threat (e.g., a measure of intelligence or academic ability) they performed significantly poorer than when the task was presented as a challenge (e.g., a potentially difficult task from which much useful skills or knowledge could be learned). In our study, the concept inventories were introduced as neither a threat nor a challenge – rather the emphasis of the exercise was placed on completion of the task. As a result, we may have created an environment that reduced the activation of stereotype threat, which could explain the lack of achievement gap between groups of students.

Finally, the changing demographic of undergraduate students across the nation may impact the stereotypes students identify, the subsequent stereotype threats they are at risk of confirming, and ultimately, their performance and persistence in science. For example, we note that the student population sampled in this study differs substantially from the population studied in Miyake *et al.* (2010), with weaker academic preparation, based on composite and subject area ACT scores and high school GPA of entering freshmen. As a result, the aspirations, motivations, and self-efficacy of students in this study may differ markedly from those students attending a more competitive school like that studied by Miyake *et al.* (2010).

Implications

Introductory science courses are diverse, complex systems with the potential to impact learning in multiple and sometimes unanticipated ways. Course context, including decisions about instructional practices, in concert with the changing demographic of our undergraduates, may reduce or enhance the prevalence of a gender achievement gap, as mediated by stereotype threat endorsement. As this research shows, gender achievement gaps are not a certainty in the science classroom, and stereotype threat endorsement may reflect factors of which we are currently unaware. We believe this research supports recent calls from the DBER community (Singer *et al.*, 2012) for replication studies that investigate the role of gender in learning undergraduate science across a variety of course settings, time, and different outcome measures.

References

Allport GW (1954). The Nature of Prejudice, Reading, MA: Addison-Wesley.

Alter AL, Aronson J, Darley JM, Rodriguez C, Ruble DN (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. J Exp Soc Psychol 46, 166–171.

Ambady N, Paik SK, Steele J, Owen-Smith A, Mitchell JP (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. J Exp Soc Psychol *40*, 401–408.

American Association of Medical Colleges (2012). MCAT scores and GPAs for applicants to US medical schools, 2002-2011. Washington, D.C.

Aronson J, Fried CB, Good C (2002). Reducing the effects of stereotype threat on african american college students by shaping theories of intelligence. J Exp Soc Psychol *38*, 113–125.

Bandura A (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev *84*, 191–215.

Brewe E, Sawtelle V, Kramer LH, O'Brien GE, Rodriguez I, Pamelá P (2010). Toward equity through participation in Modeling Instruction in introductory university physics. Phys Rev ST Phys Educ Res *6*, 010106.

Brogt E, Sabers D, Prather EE, Deming GL, Hufnagel B, Slater TF (2007). Analysis of the astronomy diagnostic test. Astro Educ Rev *6*, 25–42.

Chase C, Chin D, Oppezzo M, Schwartz D (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. J Sci Educ Technol *18*, 334–352.

Cohen GL, Garcia J, Apfel N, Master A (2006). Reducing the racial achievement gap: a social-psychological intervention. Science *313*, 1307–1310.

Cohen GL, Garcia J, Purdie-Vaughns V, Apfel N, Brzustoski P (2009). Recursive processes in self-affirmation: intervening to close the minority achievement gap. Science *324*, 400–403.

Coletta VP, Phillips JA (2005). Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. Am J Phys *73*, 1172–1182.

Delisle M-N, Guay F, Senécal C, Larose S (2009). Predicting stereotype endorsement and academic motivation in women in science programs: A longitudinal model. Learn Individ Differ *19*, 468–475.

Dietz RD, Pearson RH, Semak MR, Willis CW, Rebello NS, Engelhardt PV, Singh C (2012). Gender bias in the force concept inventory? In: AIP Conference Proceedings. Presented at the 2011 Physics Education Research Conference, pp. 171–174.

Docktor J, Heller K, Henderson C, Sabella M, Hsu L (2008). Gender differences in both force concept inventory and introductory physics performance. In: AIP Conference Proceedings. Presented at the 2008 Physics Education Research Conference, AIP, pp. 15–18.

Guiso L, Monte F, Sapienza P, Zingales L (2008). Culture, Gender, and Math. Science *320*, 1164–1165.

Hake RR (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. Am J Phys *66*, 64–74.

Hedges LV, Friedman L (1993). Gender differences in variability in intellectual abilities: A reanalysis of feingold’s results. Rev Educ Res *63*, 94–105.

Hedges LV, Nowell A (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. Science *269*, 41–45.

Hewlett SA, Buck Luce C, Servon L, Sherbin L, Shiller P, Sosnovich E, Sumberg K (2008). The Athena Factor: Reversing the Brain Drain in Science, Engineering, and Technology. Harvard Business School.

Hill C, Corbett C, St. Rose A (2010). Why So Few? Women in Science, Technology, Engineering, and Mathematics. American Association of University Women. Washington, DC.

Hufnagel B, Deming GL, Landato JM, Hodari AK (2004). The effect of stereotype threat on undergradautes in an instroductory astronomy class. J Women Minor Sci Eng *10*, 89–98.

Hyde JS (2005). The gender similarities hypothesis. Am Psychol *60*, 581–592.

Hyde JS, Lindberg SM, Linn MC, Ellis AB, Williams CC (2008). Gender similarities characterize math performance. Science *321*, 494–495.

Johns M, Schmader T, Martens A (2005). Knowing is half the battle teaching stereotype threat as a means of improving women's math performance. Psychol Sci *16*, 175–179.

Kane JM, Mertz JE (2012). Debunking myths about gender and mathematics performance. Notice AMS *59*, 10–21.

Kost LE, Pollock SJ, Finkelstein ND (2009). Characterizing the gender gap in introductory physics. Phys Rev ST Phys Educ Res *5*, 010101.

Kost-Smith LE, Pollock SJ, Finkelstein ND (2010). Gender disparities in second-semester college physics: The incremental effects of a "smog of bias". Phys Rev ST Phys Educ Res *6*, 020112.

Kray LJ, Thompson L, Galinsky A (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. J Pers Soc Psychol *80*, 942–958.

Kreutzer K, Boudreaux A (2012). Preliminary investigation of instructor effects on gender gap in introductory physics. Phys Rev ST Phys Educ Res *8*, 010120.

Langer EJ, Bashner RS, Chanowitz B (1985). Decreasing prejudice by increasing discrimination. J Pers Soc Psychol *49*, 113–120.

Locksley A, Borgida E, Brekke N, Hepburn C (1980). Sex stereotypes and social judgment. J Pers Soc Psychol *39*, 821–831.

Lorenzo M, Crouch CH, Mazur E (2006). Reducing the gender gap in the physics classroom. Am J Phys *74*, 118.

Martens A, Johns M, Greenberg J, Schimel J (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. J Exp Soc Psychol *42*, 236–243.

Marx DM, Roman JS (2002). Female Role Models: Protecting Women's Math Test Performance. Pers Soc Psychol Bull *28*, 1183–1193.

Marx DM, Stapel DA (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. J Pers Soc Psychol *91*, 243–254.

McCullough L (2002). Women in physics: A review. Phys Teach *40*, 86–91.

McCullough L, Meltzer DE (2001). Differences in male/female response patterns on alternative-format versions of FCI items. In: AIP Conference Proceedings. Presented at the Physics Education Research Conference, pp. 103-106.

McIntyre RB, Lord CG, Gresky DM, Frye GDJ, Bond Jr CF (2004). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. Cur Res Soc Psychol *10*, 116–136.

Miyake A, Kost-Smith LE, Finkelstein ND, Pollock SJ, Cohen GL, Ito TA (2010). Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation. Science *330*, 1234–1237.

Moss-Racusin, CA, Dovidio, JF, Brescoll, VL, Graham, MJ, Handelsman, J (2012). Science faculty's subtle gender biases favor male students. PNAS *109*, 16474-16479.

Mulvey PJ, Nicholson S (2008). Enrollments and degree report, 2006 (No. R-151.43). American Institute of Physics Statistical Research Center.

National Science Foundation, Division of Science Resources Statistics, 2011. Women, minorities, and persons with disabilities in science and engineering: 2011 (No. Special Report NSF 11-309). Arlington, VA.

Niederle M, Vesterlund L (2011). Gender and Competition. Annu Rev Econ *3*, 601–630.

Nosek BA, Smyth FL, Sriram N, Lindner NM, Devos T, Ayala A, Bar-Anan Y, Bergh R, Cai H, Gonsalkorale K, Kesebir S, Maliszewski N, Neto F, Olli E, Park J, Schnabel K, Shiomura K, Tulbure BT, Wiers RW, Somogyi M, Akrami N, Ekehammar B, Vianello M, Banaji MR, Greenwald AG (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. PNAS *106*, 10593–10597.

Pollock SJ, Finkelstein ND, Kost LE (2007). Reducing the gender gap in the physics classroom: How sufficient is interactive engagement? Phys Rev ST Phys Educ Res *3*, 010107.

Schmader T, Johns M, Forbes C (2008). An integrated process model of stereotype threat effects on performance. Psychol Rev *115*, 336–356.

Schwartz DL, Martin T (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. Cognition Instruct *22*, 129–184.

Simard E, Henderson AD, Gilmartin SK, Schiebinger L, Whitney T (2008). Climbing the technical ladder: Obstacles and solutions for mid-level women in technology. Michelle R. Clayman Institute for Gender Research, Stanford University, & Anita Borg Institute for Women and Technology, Stanford, CA.

Singer SR, Nielsen NR, Schweingruber HA (2012). Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering. Washington, DC, The National Academies Press.

Spencer SJ, Steele M, Quinn DM (1999). Stereotype threat and women's math performance. J Exp Soc Psychol *35*, 4–28.

Steele CM (1997). A threat in the air: How stereotypes shape intellectual identity and performance. Am Psychol *52*, 613–629.

Steele CM, Aronson J (1995). Stereotype threat and the intellectual test performance of African Americans. J Pers Soc Psychol *69*, 797–811.

Steele CM, Spencer SJ, Aronson J (2002). Contending with group image: The psychology of stereotype and social identity threat. In: Advances in Experimental Social Psychology. Academic Press, pp. 379–440.

Stoet G, Geary DC (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? Rev Gen Psychol *16*, 93–102.

Walton GM, Cohen GL (2007). A question of belonging: race, social fit, and achievement. J Pers Soc Psychol *92*, 82–96.

Walton GM, Cohen GL (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. Science *331*, 1447–1451.

Willoughby SD, Metz A (2009). Exploring gender differences with different gain calculations in astronomy and biology. Am J Phys *77*, 651.

Yeager DS, Walton GM (2011). Social-psychological interventions in education: They're not magic. Rev Educ Res *81*, 267–301.

CHAPTER 2. STEREOTYPE THREAT TRIGGERS AND CLASSROOM PRACTICES

Introduction

What's the problem?

Every year, more women leave the fields of science, technology, engineering, and mathematics (STEM) and opt for other avenues of undergraduate majors, graduate programs, research and professional vocations (National Science Foundation, 2011). This exodus away from STEM fields has been referred to in the literature as "the leaky pipeline". While certainly not the only cause of the attrition of women out of science, stereotype threat has been indicated as one of the main factors in a recent meta-analysis (Hill et al., 2010).

The goals of this paper are to:

1. Provide a recent meta-analysis of selected literature regarding stereotype threat and its implications for threatened groups. This particular area is saturated with empirical research, and the following literature review should be considered only a small glimpse of particularly relevant work;
2. Describe probable ways that stereotype threat impairs performance both physiologically and cognitively;
3. Identify and classify triggers of stereotype threat, specifically, stereotype threat that affects women in STEM. Women bring invaluable and unique perspectives and skill sets to STEM fields; as a result, many women are actively recruited to STEM fields, yet may not persist in these fields, in part because of the obstacle stereotype threat has placed in the way. Identifying the triggers to stereotype threat activation in the classrooms leading up to major life and work decisions may keep more women on track to be higher-ranking and higher-paid individuals in the STEM workforce; and,

4. Recommend best practices for reducing or eliminating stereotype threat in the undergraduate classroom.

In order to prepare a diverse and productive STEM workforce, educators must work especially hard to ensure each student at the undergraduate level is getting the most out of his or her college education and experience. Identifying the triggers of stereotype threat may change the outlook for women and minorities in STEM careers. With tuition rates increasing approximately 1,120% within the last 30 years, student debt rates outrivaling car loans and credit cards as the single largest sources of personal debt approaching the tune of $1 billion, and falling returns on educational investment within a poor economy, students today are struggling more than ever before (dailyfinance.com). Educators at all levels have an obligation to ensure equal access to all majors, especially STEM majors, and an obligation to provide an education that reflects the current occupational needs of society.

Methods

This particular topic of research is laden with social psychological terminology in which many readers may be unfamiliar. Table 8, below, familiarizes readers with some of the central concepts revolving around stereotype threat and its triggers. I realize that some of these concepts and terminology may have other meanings in different contexts, but for this paper, these definitions are best suited.

As stated earlier, the literature included in this meta-analysis is just a small glimpse of that which has been studied in the last 20 years. This area has been so extensively studied that if one were to input “stereotype threat” into a large research database, such as Google Scholar, it would yield upwards of 12,400 returns. Specifically searching since the landmark paper by Steele and Aronson in 1995, 11,500 papers are returned. This amount of literature is exhaustive and required a focusing of the search parameters. Building upon Steele and Aronson (1995), additional searches were done using the keywords “undergraduate STEM” and “gender” in addition to “stereotype threat”. From these keywords, 46 publications, and 8,820 publications were returned, respectively. These papers were sorted through by most relevant content, while additional papers were included if they were cited by papers who initially cited Steele and Aronson, 1995 or Steele 1997.

Table 8. **Important terminology related to stereotype threat and accompanying definitions.**

Terminology	Definition
Domain	A location or situational realm in which an individual is expected to implement either a social or personal identity. Examples include: School, family, race, culture, dance class, etc.
Domain Identification	Cognitively defining one’s self-concept or self-definition to be aligned with good prospects or positive outcomes within the domain (Steele & Aronson, 1995).
Gender Achievement Gap	A disparity between the educational achievement of males and females in favor of one gender over the other.
Gender Bias (unconscious)	The tendency to unconsciously prefer one gender to another.
In-Group	Collections of two or more people similar to the self, sharing the same or similar social identity (Turner, 1982), e.g., “We” or “Us”.
Out-Group	Collections of people dissimilar to the self who do not share the same or similar social identities (Turner, 1982), e.g., “Them”.
Stereotype Threat	The situational predicament in which an individual is at risk for confirming a negative stereotype about his/her social group, based on performance in a given situation (Steele & Aronson, 1995).
Stereotype Threat Endorsement	To believe in or support a particular stereotype as being reflective or true of oneself, or an in- or out-group member or their abilities.
Threatening Environment	Any environment in which an individual’s performance on a task can be viewed with a negative connotation (Inzlicht & Schmader, 2012).

Discussion

Why do women leave?

Women leave STEM fields for a variety of reasons, but here I focus on those related to stereotypes of women in STEM. Biases toward achievement capabilities of women in STEM (e.g., women are not as good at math and science as their male counterparts) are pervasive in STEM. Recent research, sampling research-intensive universities, has demonstrated that an unconscious gender bias against undergraduate women exists among science faculty members across the country, regardless of their own personal gender (Moss-Racusin, Dovidio, Brescoll, Graham & Handelsman, 2012). This bias may contribute to gender disparity in academic science programs.

One of the reasons many women leave STEM fields also includes, as previously indicated, stereotype threat, a situationally-dependent, social and psychological phenomenon. Although many students pick a major based on factors like time commitment, costs, department reputation, location, and student morale, for a stereotyped individual, an additional factor may contribute to this decision: perceived stereotype threat. For stereotype threatened individuals, the fear of frequent activation may play a more active role in the decision to enter in STEM fields, either as a college major or when joining the workforce.

In the formal education environment, stereotype threatened individuals must overcome assessment bias. Standardized exams, for example, usually measure intellectual performance via comparison to another group. It is often difficult for educators, college entrance officials, and future employers to not equate inferior performance with inferior ability (Howard & Hammond, 1985; Walton & Spencer, 2009). Educators, college entrance committees, and future employers have preconceived notions about what a successful candidate may be able to achieve as a result of standardized exam scores. For example, on the 2008 Math portion of the SAT, stereotype threat could have accounted for a deficit of 19-21 points for those women threatened (Walton & Spencer, 2009) due to statistical analysis of latent student achievement in multiple studies of documented achievement gaps. The mean gender achievement gap between males and females on this exam was 30 points, making the 20-point deficit even more crucial to closing the achievement gap (Hill et al., 2010).

What is stereotype threat?

Stereotype threat is a social-psychological threat in which participants in a given situation are at risk of confirming an existing negative stereotype about a group with which the participants identify (Steele, 1997). Past research has documented that achievement of African Americans (Cohen, Garcia, Apfel, & Master, 2006; Steele & Aaronson, 1995), women (Hill et al., 2010; Spencer, Steele, & Quinn, 1999), Hispanics and other minorities (Gonzales, Blanton, & Williams, 2002; Schmader & Johns, 2003; Aronson & Salinas, 1997), individuals with low socioeconomic status (Croziet & Claire, 1998; Spencer & Castano, 2007) and even white men (Aaronson, Lustina, Good & Keough, 1999, Koenig & Eagly, 2005) on different exams and tasks is vulnerable to stereotype threat if the individual identifies with the domain and is aware of a negative stereotype about their group. In order for these individuals to be domain-identified, they need to perceive success within the domain first (i.e., have the resources and skills required to theoretically advance and succeed), while possessing the personal sense of being acknowledged and valued in the domain (Steele, 1997).

Anyone who is identified with a domain feels the urge to belong, achieve, identify and participate within that particular domain and its subdomains (Steele, 1997). As an example, for a woman to be domain identified with math, she must perceive success within math, and feel the urge to belong in math and math-related subdomains like physics and engineering. These domain-identified students are thought to be the ones who are the most at risk for stereotype threat endorsement. For a negative stereotype to evoke threat, the domain-identified students must feel the threat is self-relevant (Steele, 1997). However, to experience stereotype threat, one may not need to subscribe in, internalize, nor feel that the threat has any impact on themself, but because they are a member of their group and they are identified within the domain in which the negative stereotype is salient, the student will feel concerns about being stereotyped in that domain (Steele, 1997). It is important to note here that every student will not feel the same level of threat, endorsement, or domain identification, which makes stereotype threat a very difficult and complex issue to study.

Where do stereotypes come from?

At a very basic level, a stereotype is a belief held by different groups about a particular person or type of people that is an oversimplified idea, evaluation, or opinion. These opinions are either learned or based on an observation and evaluated in a simplified way; as a result, using the lens of attitude formation is useful. Attitudes are cognitive impressions, such as feelings or beliefs, that guide our reactions to and thoughts about people, things, or events (Myers, 2004) and attitude formation, therefore, would be the act of converting a feeling or belief into a positive or negative evaluation.

Gawronski and Bodenhausen (2006) argue there are two different types of attitude and two underlying types of mental processes important for formulating each type of attitude. The first are explicit attitudes, which are deliberate attitudes, recently acquired, and more likely to change. Explicit attitudes are created by propositional reasoning, which is based on factual validation of evaluations and beliefs (Gawronski & Bodenhausen, 2006). These types of attitudes are usually uncovered in research studies via self-report measurements. In contrast, the second type of attitude is implicit attitude, which are attitudes created automatically “on the spot”, moderated by underlying mental associative processes, and thought to be more robust than explicit attitudes (Gawronski & Bodenhausen, 2006). These associative processes are activated independent of one's opinion of truth or falsity (Gawronski & Bodenhausen, 2006). Implicit attitudes are more stable than explicit attitudes, and are deeply rooted as a result of long-term socialization processes (Gawronski & Bodenhausen, 2006.) Implicit attitudes are usually measured in terms of Implicit Association Tests, whereby a reaction time is measured when associating two different groups of words, for example, African Americans and athletic ability versus Whites and athletic ability. Explicit and implicit attitudes, and their underlying mental processes, can work together or separately to make evaluations; however, I believe that implicit attitudes are a more likely source than explicit attitudes when it comes to creating stereotypes. Once stereotypes are formed, they often have the ability to spread to others. Eventually, they become pervasive and have the ability to threaten individuals in social situations, including the classroom.

Why does stereotype threat occur?

Answering the question of why stereotype threat occurs is not easy. It may reflect attitude formation, social identity, ability and preparation, or a complex combination of interactions among all three.

Experimental evidence of negative attitudes of African American subjects found the subjects to have very strong in-group bias, or preferential treatment toward those perceived to be within one's group versus those who are not, in situations of perceived negativity of explicit attitude formation (Livingston, 2002). In one experiment, perceived negativity was calculated by administering a racial version of the Collective Self-Esteem Survey (Crocker et al., 1994), which measures explicit attitudes (deliberate) of another group's appraisal of an in-group. African Americans show no in-group bias, and even out-group favoritism when they make implicit (automatic) attitude evaluations in the same situations (Livingston, 2002a). Further results of this research article show that African Americans with more contact with whites, versus those with less contact, tend to more accurately assess white's perceived negativity towards their black in-group, which is experimentally accurate of the statistical ratings found of whites (Livingston, 2002b). This study signaled that implicit attitude formation, that which is thought to be more robust and a result of socialization experiences (Gawronski & Bodenhausen, 2006) of the in-group, is vulnerable to perceived negativity. To transfer this information to an example, women who are in math classes are likely aware of the perceived negativity of the out-group (i.e., white males), and could possibly accurately assess the degree of the negativity, and are yet subjecting themselves to the same environment week after week.

Implicit attitudes about the subject of mathematics have also been measured in K12 and college students (Nosek et al, 2009; Nosek, Banaji, & Greenwald, 2002). Implicit attitudes were measured with an Implicit Association Test in which participants were given two choices of concepts, one "math" and the contrasting category "arts". A participant had to select which of the two concepts with which the participant more easily associated while being timed. Women demonstrated more perceived negativity, that is responding faster to computer items involving arbitrary letters faster than toward a string of arbitrary numbers, more quickly associating the woman/math relationship as unpleasant, and by more quickly selecting a random unfavorable geographic location as more pleasant than math (Nosek, Banaji, & Greenwald, 2002). In an additional study, men also showed a perceived negativity toward math and science in comparison with the arts and language, but less so than females. Male students had more positive implicit and explicit attitudes toward the gender stereotype Math = Male, than did women (Nosek, Banaji, & Greenwald, 2002.) Even at the national level, a bias toward gender-science stereotyping in the United States could be an attributing factor in gender differences in achievement of 8th-graders, (Nosek et al, 2009), underscoring the importance of educators at all levels to be aware of stereotype threat, and the large impacts this social-psychological threat can play in society.

Another explanation for why stereotype threat occurs emerges from Social Identity Theory (Tajfel, 1981). Toni Schmader, in 2002, used the framework put forth by Tajfel in a novel way in order to recognize why stereotype threat occurs. The integrated process model of stereotype threat effects on performance (Schmader, Johns, and Forbes, 2008) is well-articulated, and has been extensively referenced in the literature. Schmader et al focused on the intergroup interactions relating to stereotype threat whilst creating this model, in addition to showing how social and psychological processes interact to interfere with working memory. The strengths of this model are numerous: it integrates diverse and far-reaching topics in the literature, it is well supported by previous literature published, it highlights the convoluted nature of stereotype threat, and it depicts both physiological and psychological processes combining, which ultimately leads to impairment. This model, however, does not include specific descriptions for triggers of stereotype threat, and also did not explicitly include social identity theory, and cognitive dissonance theory, two very important social psychological theories with major impacts on stereotyped individuals. I extend this model to include the three types of triggers of stereotype threat, social identity theory and cognitive dissonance theory (Figure 3), which together, more fully capture how stereotype threat is activated and contributes to the leaky pipeline. Each of these additions are of import due to their ability to further explain how the monitoring process is heightened in situations of stereotype threat, and as an explanatory tool for elucidating why women and minorities ultimately disengage or permanently leave the threatening situation, which will be explained later.

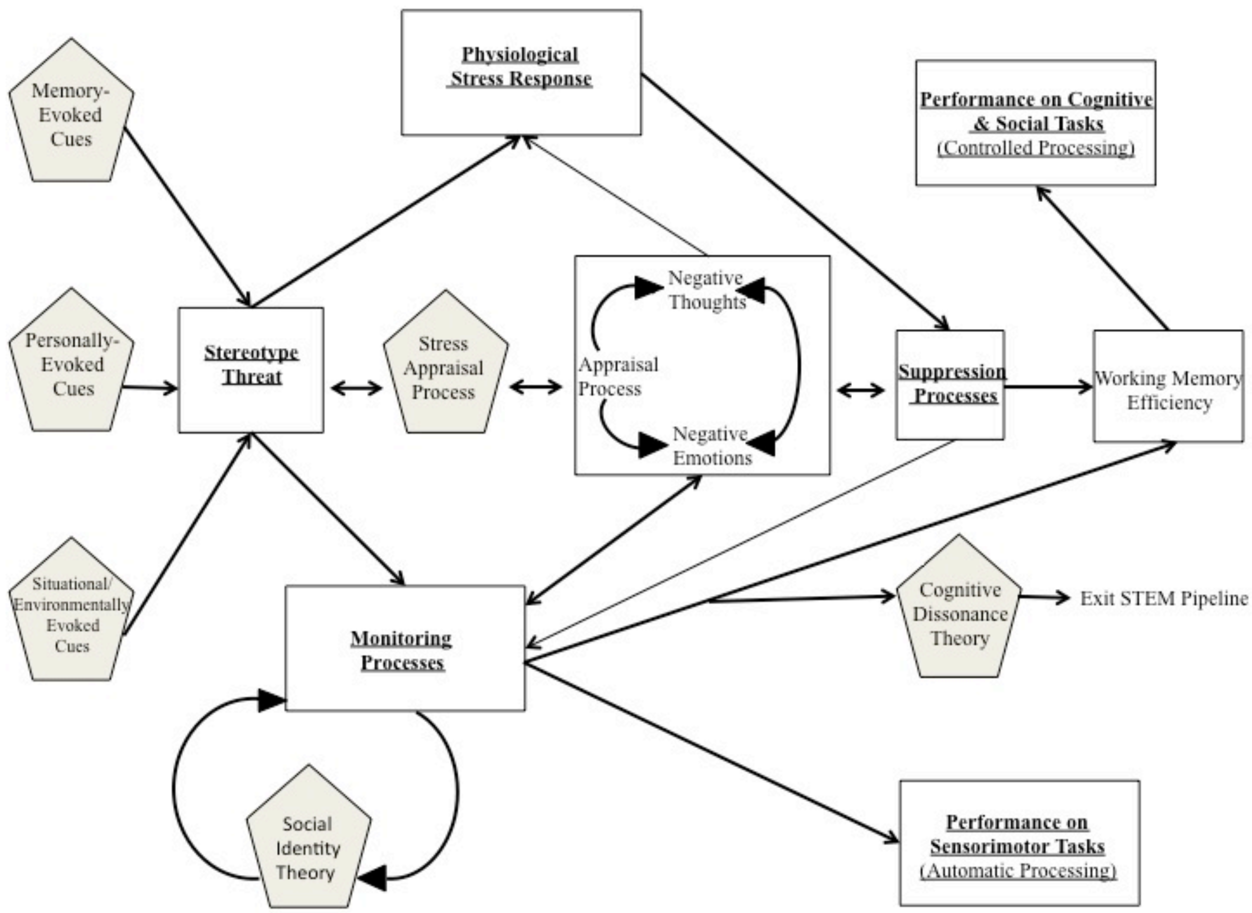

Figure 3. The original Schmader, Johns, and Forbes (2008) Integrated Process Model of Stereotype Threat effects on performance, with social identity theory and cognitive dissonance theory added to show additional effects on working memory and a possible explanation for the "leaky pipeline" of women in STEM fields.

Social Identity Theory is very important for intergroup relations. It is an analysis of one's individual role, known as personal identity, and the individual's multiple social roles, known as social identities (Tajfel & Turner, 1986). Examples of personal identities could be characteristics of an individual's personality that describe the individual "I", whereas examples of social identities could be religion, ethnicity, culture or sex, "we" or "us" (Hogg, 2006). The way social identity theory plays into stereotype threat is that for each individual social situation, a person has control over the one salient psychological identity they are committed to at the time (Hogg, 2006). Since each person can only have one salient identity at a time, he or she may switch social identities depending on the need of the situation. In the revised Schmader model (Figure 3), during the monitoring processes, a recursive loop is added which is drawn backward to Social Identity Theory. This loop indicates the social identity being “tried on” and “taken back off” and switched for a new, more appropriate identity within a threatening environment or a new threatening group. By utilizing this heightened monitoring process, an individual may be able to select the right social identity and not experience stereotype threat, or they may select the wrong one, and fall victim.

In general, people are motivated to keep both their personal and social identities positive. One’s personal identity can be less positive if compared to another individual better at a given task, and one’s social identity can be less positive when compared to any out-group in an unfavorable light. (e.g., Ray is better at math than Jessica; men are better at math than women). If there exists a negative stereotype that acts upon one of their various social identities, and it leads them to compare themselves unfavorably to the out-group, the person will experience not just a less positive identity, but threat (Schmader, 2002). Empirical support for this theory was found in an experiment demonstrating that for women identifying strongly with their gender, the stereotype that women are not as good at math as men interferes more so than a group of women identifying less strongly with their gender; their subsequent ability to perform well on an exam is inhibited due to their strong female identification, which is likely interfering with their working memory efficiency (Schmaeder, 2002). Much like domain identification (Steele, 1997), mentioned above, this work provides further evidence that group identification is also an important component to activating stereotype threat.

Stereotype threat at the college level could also occur as a result of freshmen students coming into a science department with differential mathematics preparation and ability levels

(Olson & Riordan, 2012). This lack of preparedness in some students may manifest as stress during the semester. In the 1960's, Richard Lazarus created an emotional framework for processing psychological stress that may help explain how a student approaches a stereotypical task or environment, called the Stress-Appraisal Process. In Figure 3, the Stress-Appraisal Process is depicted, moving between stereotype threat and the appraisal processes box. When faced with a new stressor, the subject makes an unconscious assessment or appraisal of his or her personal resources that should help him/her either overcome the stressor or not (i.e., lots of math background, confidence, supportive family members, past instructors, etc., or lack thereof) (Lazarus, 1999). If the student feels that they do have the personal resources to overcome the stressor, they view it as a challenge, whereas if they view the demands of the stressor as greater than their resources, they view the situation as threatening, inflicting negative emotions and causing a strong response (Lazarus, 1999), sometimes physiologically.

How does stereotype threat impact performance?

Delisle, Guay, Senecal, and Larose (2009) state three postulates that need to occur in order for stereotype threat to affect the performance of an individual: (1) there needs to be a known stereotype about a social group revolving around achievement in a particular field, (2) the individual needs to identify with the field in question, and have importance attached to it, and lastly, (3) the individual must be faced with a task that ultimately could reveal the stereotype (2009). Additional effects of stereotype threat activation can include decreased autonomic academic motivation or motivation resulting from positive or pleasurable feelings about the field (Delisle, Guay, Senecal, & Larose, 2009), along with decreased domain identification (Steele, 1997), and leaving the domain altogether in an effort to protect self-esteem (Aronson, Fried, & Good, 2002; Aronson, Quinn, & Spencer, 1998; Steele, Spencer, Aronson, 2002). After a brief review of physiological impacts of stereotype threat (see below), I will use Cognitive Dissonance

Theory (Festinger 1957) to explain how stereotype threat poses negative cognitive impacts upon it's target, and how this leads qualified women to eventually leak out of science and math.

How stereotype threat impacts physiology

Work by Schmader, Johns, & Forbes (2008) suggest stereotype threat can inhibit performance by three interrelated cognitive mechanisms: (1) stressful tasks increase the body's physiological response, decreasing the processing ability of the prefrontal cortex, (2) stressful tasks create a tendency for sufferers to spend energy and time to actively monitor their performance, where they otherwise would not, and (3) subjects try to reduce negative thoughts and emotions, and spend needed energy on self-regulation. All three of these mechanisms work together to impair performance by utilizing cognitive energy for processes other than exam performance (Schmader, Johns, Forbes, 2008).

Experimental research has also found additional physiological responses in subjects in stereotype threatening environments, such as increased self-consciousness and reduced cognitive processing efficiency (Steele & Aronson, 1995), higher blood pressure in African Americans than in control or European Americans (Blascovich, Spencer, Quinn, & Steele, 2001), reduced working memory capacity in women and Latinos who are the target of a primed stereotype (Schmader, & Johns, 2003), and decreased heart rate variability of students during an ability-diagnostic exam (Croizet, Despres, Gauzins, Huguet, Leyens, & Meot, 2004). These findings corroborate work by O'Brien and Crandall in 2003, postulating that theories of arousal (i.e., heightened sympathetic nervous system activity: sweating, increases in blood pressure, stress hormone release, etc.) are consistent with stereotype threat effects. Arousal theory, sometimes referred to as the Yerkes-Dodson Law, states that every task has an optimal amount of arousal required to accomplish it optimally (Yerkes & Dodson, 1908). Tasks usually requiring fine motor skills and high amounts of concentration (i.e., difficult academic exams) are usually optimized when arousal states are low. Well-practiced tasks and those requiring lots of motor function (i.e., skateboarding, gymnastics, etc.) are optimized when arousal states are high. Arousal theory postulates that arousal should decrease achievement on difficult exams, while increasing achievement on easy exams because stress hormones released during high arousal impair cognitive functioning, leading test-takers to select the wrong choices and make more mistakes. For example, women who took a math exam and were informed that women underperform men on this exam (creating stereotype threat and physiological arousal) did worse than women who were blind to the differences (O'Brien & Crandall, 2003).

How stereotype threat affects cognition

Cognitive Dissonance Theory, developed by Leon Festinger in 1957, could be an underlying or explanatory mechanism for why stereotype threat might reduce student's performance and eventually cause them to leave the field with which they are identified. Any person experiencing cognitive dissonance (inconsistency) experiences considerable discomfort in their daily lives and therefore should be motivated to reduce the dissonance and strive for consonance. When present, dissonance can be strong enough to make individuals avoid situations and information that increase their dissonance (Festinger, 1957). As shown in Figure 3, dissonance is also strong enough to hijack working memory efficiency. Dissonance then increases and festers, and can ultimately be a reason why women decide to leave STEM disciplines permanently.

Let's use, for an example, a woman who is in a physics class. She is aware of the stereotype that men perform better at mathematics and science than women do. She is also aware that she is a woman and a minority in the classroom. This combination of factors causes her dissonance. She can do one of two things in this situation. She can take steps to reduce her dissonance by quitting the physics class or she can change her knowledge by reading literature about important female figures in the field of physics, thus reinforcing the positive image that she is a smart woman and can do math and science as well as or better than anyone else. Festinger believed that once dissonance is created, it will persist and efforts to reduce it will not cease, as people always gravitate toward and seek consonance. Future efforts to reduce this dissonance could include switching majors, leaving her career, and trying harder to prove herself. Studies show women have a tendency to work harder in their fields by taking on extra shifts and working more hours (Lemkau, 1979), possibly to rectify dissonance. The dissonance remains, so a woman might add a new cognitive element, for example, seeking out a female mentor. During this time, she may also decrease addition of new cognitive elements by staying away from glass ceiling and gender inequality literature, which would make the cognitive dissonance stronger. These strategies for reducing dissonance and deferring stereotype threat activation may help a woman make it through the physics class, and she may receive a high grade, but Festinger states that she will never be able to guarantee that dissonance will be completely reduced if she continues to be domain identified in a field where stereotypes create dissonance. This cognitive frame could be a mechanism to explain the "the leaky pipeline" and its persistence despite years of active recruitment of women into STEM fields (Hewlett, Buck Luce, Servon, Sherbin, Shiller, Sosnovich, & Suremberg, 2008).

Personally dealing with stereotype threat is difficult and proving oneself in one domain, like a physics class one semester, is not an exemption to proving oneself in another, very similar domain (i.e., a more difficult physics class the next semester). A side effect of proving oneself time and time again is that women start developing goals for their short-term success based on their abilities, but if their abilities fail, they blame themselves and become discouraged (Grant & Dweck, 2003). The higher women go in the domain of math and science, the more likely they will be the minority (Steele, 1997). In order to prevent this consequence, and others, identification of stereotype threat triggers may help ameliorate the achievement gaps between minorities and majorities, giving aspiring underrepresented groups a more level playing field.

What may trigger stereotype threat?

It is important to note that some students, for whatever reason, are not affected by stereotype threat. The exact reason is often unknown and it can be difficult to explain why some students are more or less susceptible to stereotype threats (Schmader, 2002). Some people differ in their amount of *stigma consciousness* - their expectations about the extent they expect to be stereotyped or discriminated by others (Pinel, 1999). Some students could be differentially less domain identified than others, some could be less gender identified than others, some could be less identified in any one of possibly hundreds of social domains, making this research area, again, a complex one in which to unravel individual differences.

While some students do not experience stereotype threat, over 300 experimental cases of stereotype threat have been published within the last 20 years (Aronson & Dee, 2012). Most of these experiments were done in the laboratory, and although less fully applicable to the "real world," are still interesting in that small, subtle, and often unnoticed cues from a situation have profound consequences. More recent research has been done to corroborate some of the findings, and have utilized an in situ approach to experimental design (Ford et al., 2004; Keller & Dauenheimer, 2003). An extensive review of the literature for research identifying triggers leading up to stereotype threat has yielded results that can be divided into 3 distinct divisions - memory-evoked cues, personally-evoked cues, and situational/environmentally-evoked cues, the latter of which is the most well-studied. Table 9 shows each of the cues, and provides an explanation of the rationale for each of the three distinct divisions.

Table 9. **Types of triggers that evoke stereotype threat, and the rationale behind each division.**

Triggers of Stereotype Threat	**Description**
Memory-Evoked Cues	Stereotype threat is retriggered by implicit memories from past stereotype threat activation (Smith & Branscombe, 1988) and can not be prevented.
Personally-Evoked Cues	Cues directed toward an individual's looks, actions, or abilities based on domain identification, rather than situations (i.e., Steele, 1997).
Situational/Environmentally-Evoked Cues	• Evoked by "threatening environments"- situational instances where a person is in the company of others who activate a negative stereotype (Delisle, Guay, Senecal, Larose, 2009). • Pose a direct threat to self-identity (Schmader, Johns, & Forbs, 2008).
Two subtypes: A. Experimental and Academic Manipulated Situations	A. Evoked by manipulating laboratory or *in situ* circumstances and have provided the largest body of evidence for stereotype threat's existence. These situations are where the majority of the data on stereotype threat come from.
B. Sheer Number of Similar Others	B. The number of fellow group members around an individual who can revoke the threat as postulated by Social Impact Theory (Latane, 1981). The pressure one source exerts on an individual decreases as a power function with the number of targets available (Latane, 1981).

Situational/environmental cues are evoked by threatening environments, providing an explanation as to why students who are susceptible to stereotype threat are not in a perpetual state of threat in every classroom throughout the day. A "threatening environment" includes situational instances where a person is in the company of others who activate a negative stereotype (Delisle, Guay, Senecal, Larose, 2009), such as a teacher stating the difficulty of a math exam in the instructions. This situation can be anywhere a person could possibly be exposed to a negative stereotype about one's group - it does not have to be an official location, such as a classroom, work place, or meeting place (Lawrence, Crocker, & Dweck, 2005). An example of an atypical situation would be at a wedding dance, where the DJ makes a comment about white men's abilities on the dance floor. Threatening situations pose a direct threat to an individual's self-identity, a set of beliefs one holds about themselves, or their own awareness that they are a unique individual with an identity, who is capable of change in differing environments (Schmader, Johns, & Forbs, 2008).

Situational/environmentally-evoked cues can be further divided into two categories. The first broad category is composed of many examples of both experimental and academic situations that are likely to evoke stereotype threat within a subject. One way to evoke a stereotype is to make group membership salient and probe a stereotyped student's ability in a threatening situation (Lawrence, Crocker, and Dweck, 2005). Similarly, one can induce threat by subjecting the participant to difficult or advanced curriculum. Work by Spencer, Steele, & Quinn (1999) found the task must be difficult enough for the participant to struggle and that struggle could be the same as one highlighted by a popular stereotype. Along with the difficulty level of the exam, the departmental structure and the conditions under which the content is taught, are also important. A lack of female role models, or an overrepresentation of male models can create a threatening environment (Marx & Roman, 2002; McIntyre, Lord, Gresky, Ten Eyck, Frye, & Bond, Jr., 2005). When female engineering students interact with sexist men, social identity threat is induced (Logel, Walton, Spencer, Iserman, 2009). If the atmosphere or assessment within the classroom under which the content is taught is highly competitive, women have been shown to perform differently than men on math-related tasks (Niederle & Vesterlund, 2011). Even non-academic general situational cues, like television commercials depicting women performing stereotypical tasks, had an affect on women's math ability on exams, causing them to avoid math problems on a combination math and verbal exam, and to eventually indicate less interest in future workforce vocations involving math and quantitative reasoning (Davies, Spencer, Quinn, & Gerhardstein, 2002).

The second category of situational/environmentally-evoked cues consists of situations involving the sheer number of fellow group members around an individual to help revoke the threat, as postulated by the Social Impact Theory (Latane, 1981). The situational cues involving the number of women in a given domain are well researched and have a foundation stemming from Social Impact Theory (Latane, 1981), which states that the pressure one source (in this case, stereotype threat) exerts on an individual target decreases as a power function with the number of targets available (in this case, the number of minorities in a given threatening situation). So, according to this theory, the more women in a math class, the less stereotype threat they each will feel. Research confirms these findings. Women underrepresented by sheer number in a group do poorer on a math exam, are assessed as having more stereotype-related anxiety, and are more likely to think about the negative gender stereotypes that apply to them, compared to when they take a math exam in a group with similar or equal representation (Beaton, Tougas, Rinfret, Huard, & Delisle, 2007; Inzlicht & Ben-Zeev, 2000). However, in some circumstances, the threat due to being outnumbered is so blatant that we see some women in the "token" position actually over perform as a way of compensating for their underrepresented numbers within a classroom (Kray et al., 2001).

The importance of equal proportions of men and women extends outside the classroom and into the workplace. In Kanter's book, *Men and Women of the Corporation* (1977), she explains how the relationship between men and women, as minorities in the workplace can be explained by breaking the representative numbers down into proportions. For example, if a particular group, like women, makes up less than 20% of the group, they are in the "token position", in which members are isolated and seeking out alliances is more difficult; women or minorities who make up 35% of the group are said to be in a "minority position". The token women feel an increased need to prove themselves to the majority, possibly to try to deflect stereotypical views upon themselves, feel less competent conducting business in the workplace, and feel like they receive positions with lower responsibility levels (Kantzer, 1977). The minority women do better, but are not fully comfortable, as relationships with fellow male peers are still somewhat uncomfortable. However, due to their increased numbers, women in the minority condition are more easily able to form friendships and alliances with other women. These findings are important, because in some science programs, the makeup of the department may produce higher gender stereotype endorsement. Low numbers of female faculty and female students, less than 20% of the departmental makeup, makes these women subject to the problems of other minorities in the "token position" described by Kantzer (Delisle, Guay, Senecal, Larose, 2009). Kanter (1977, 1989) advocates for a more "balanced" proportion of men and women in the workplace, as it lowers threat, and takes pressure away from in-groups and focuses more closely on individuality instead. One way of reducing stereotype threat among women and increasing their test scores at the same time is to increase the visibility and interaction with strong female role models and instructors in the domain they identify with (McIntyre, Paulson, Lord, 2003; McIntyre, Lord, Gresky, Ten Eyck, Frye, Bond, Jr., 2005). For example, when a competent woman administered a math examination, the achievement gap between women and men was reduced (Marx & Roman, 2002).

One way to evoke stereotype threat that falls within both the general situational category and the personal category, described below, is to have subjects indicate their gender or race on a survey item before (Steele & Aronson, 1995) as opposed to after a difficult exam. These findings highlight the fact that some inadvertent actions of educators have a hidden and lasting impact on student achievement.

Personally-evoked cues are those that are directed toward an individual's looks, actions, or abilities, rather than situations. For example, in order for anyone to experience stereotype threat, s/he needs to be domain-identified (Steele, 1997). As discussed previously, domain identification varies greatly from person to person, but is a major component for stereotype threat to act upon. Asking students to indicate their race before an exam, to actively cue a personal trait, regardless of diagnostic ability, has a negative affect upon these students' performance (Steele & Aronson, 1995).

Lastly, stereotype threat can be re-triggered via implicit memories from past stereotype threat activation (Smith & Branscombe, 1988); therefore, anyone who has experienced threat in the past may be reinitiated to it when returned to a similar situation with a similar makeup of peers. Even possessing the knowledge and remembering that there is a negative stereotype about an identified in-groups' lower academic ability, is enough to reduce African American's achievement on academic exams (Steele & Aronson, 1995). Though these stereotypes are not primed or reminded prior, these students are still at risk due to memory of past experiences.

What should we do in the classroom?

How can triggers be overcome? By identifying and being cognizant of triggers and also implementing the best practices possible in our classrooms to reduce stereotype threat for all of our students, we will fulfill an important portion of our duty to provide an equal opportunity for students to succeed in the classroom. Table 10 is an abbreviated list of the recommendations to overcome triggers in the classroom, as identified by a review of empirical research. Remember, an instructor should not automatically assume there is a gender achievement gap within their classrooms (Lauer, et al, 2013), but it is always a good idea to prevent stereotype threat from occurring, as each classroom and the stereotype threat that could reside within is uniquely context-dependent.

Table 10. **Best practice recommendations for decreasing stereotype threat in the classroom as emphasized by peer-reviewed literature.**

Recommendations for the Prevention of Stereotype Threat in Classroom	Citation
Educate students about stereotype threat and attribute test anxiety accordingly. Reframe word choice from "math exam" to "problem-solving task".	Johns, Schmader, & Martens, 2005 Johns, Inzlicht, & Schmader, 2009
Change assessment context changed to one designed to reduce stereotype threat.	Walton & Spencer, 2009
Allow threatened students to disengage/ distance themselves from their group and stand alone as individuals.	Locksley, Borgida, Brekke & Hepburn, 1980 Langer, Bashner, & Chanowitz, 1985 Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004
Institute wise schooling can decrease threat and increase domain belongingness.	Kreutzer & Boudreaux, 2012
Revamp the course to include research-validated, highly interactive instructional techniques.	Freeman et. al., 2007 Haak et. al, 2011
Prime positive social or individual identities instead of negative ones.	Schmader, 2002
Administer values-affirmation tasks.	Miyake et al., 2010
Reframe threat as a challenge to women.	Alter, Aronson, Darley, & Rodriquez, 2010
Encourage coping sense of humor in math classroom.	Ford, Ferguson, Brooks, Hagadone, 2004
Educate students that intelligence is malleable.	Aronson, Fried, & Good, 2001

Familiarizing oneself to at-risk populations enrolled within the classroom and teaching them about stereotype threat, while attributing any anxiety felt before an exam to the threat can positively impact performance by women on math exams (Johns, Schmader, & Martens, 2005). In fact, this study also found that even a re-framing of word choice improves women's math exam scores (e.g., instead of "math exam", use "problem-solving task"). Other encouraging results for fixing this gap were realized in a related large meta-analysis also conducted by Walton and Spencer in 2009, showing that women and minorities performed better than non-stereotyped students of equal math ability when the assessment context was changed to one designed to reduce psychological threat. The meta-analysis consisted of stereotype threat intervention experiments with manipulations designed to reduce threat in the everyday classroom environment, and the measures of this post-intervention treatment performance were not other exams given in the lab, but actual classroom performance (Walton & Spencer, 2009).

Another way to reduce stereotype threat, after it has been discussed with the class, is to allow any students who are at risk for stereotype threat to separate themselves as an unique member outside of their stereotyped group (Locksley, Borgida, Brekke & Hepburn, 1980; Langer et al., 1985). Locksley et al. (1980) found that students were less stereotypical in their judgments of other students if they had contrary diagnostic or behavioral information about the displayed student's abilities. Allowing students to showcase what they have achieved in the past may be a fairly simple way for them to promote achievement in the future. This practice may allow them to distance themselves from that stereotyped group, if they prefer, while still fully participating in the class (Ambady et. al., 2004).

Though it may require the most work and possibly an extensive overhaul of the classroom, the following practices will help students feel like equals within the classroom.

Instituting wise-schooling, or a set of instructional practices borrowed from educational psychology, can decrease stereotype threat and help increase domain belongingness (Kreutzer & Boudreaux, 2012), both of which can be impactful to students. Some of these methods are easy to implement, such as instilling confidence within students by undermining prior poor performance or memorizing student names and using them frequently in the classroom, and can help impact stereotype threat by creating a more solid teacher/student relationship. (Kreutzer & Boudreaux, 2012). Some of the other practices of wise-schooling are more involved, for example, encouraging women throughout the semester that they are an important part of the physics classroom and research arena may overcome stereotype threat by affirming domain belongingness (Kreutzer & Boudreaux, 2012).

Another way to reduce stereotype threat is to highlight it in the curriculum and give students the power to deal with and overcome it. Teaching students about stereotype threat can help to increase their cognitive executive functioning on tasks (Johns, Inzlicht, & Schmaeder, 2009). By giving students the knowledge that what they are feeling and experiencing is normal, educators give these stereotyped students the opportunity to emotionally reconstruct their reactions and cognitively reevaluate a threatening situation (Johns, Inzlicht, & Schmaeder, 2009). This means that the threatened student will no longer be taxing cognitive resources, and can allocate those resources instead to the task at hand, while overcoming the physiological triggers mentioned earlier. Lastly, revamping the course to include research-validated, highly interactive instructional techniques helps close the achievement gap (Freeman et. al, 2007; Haak et. al, 2011) and decreases the amount of stereotype threat experienced by students. This practice removes some of the situational triggers, like not making mention of stereotypes or achievement gaps in the past, helps with domain identification, and does not draw attention to the sheer number of minorities in the class.

When it comes time for assessment, an instructor can decrease stereotype threat by manipulating test-taking instructions to (1) clearly indicate that the outcome of the exam as non-diagnostic of the student's abilities (Steele & Aaronson, 1995; Johns, et al., 2005), (2) reassure students that exams are fair (Good, et al., 2003), (3) reiterate that intelligence is malleable and not fixed, (4) reassure students that no gender achievement gap exists among test takers of different groups, and (5) that the exam will not contain any cultural biases, because the meaning of the exam may be altered in the minds of some stereotyped students as a result (Lawrence, Crocker, & Dweck, 2005). It is also a good idea during the assessment cycle to provide critical feedback as it has been shown to be a strong motivator of African American students when the instructor praises their successes and future potential (Cohen, Steele, & Ross, 1997).

Finally, some advice for the day-to-day classroom. A teacher can prime positive identities such as "Asian" instead of "woman" before a difficult math exam or homework problem (Schmader, 2002). As an instructor, it is possible to reduce gender gaps by not mentioning the existence of any gender gaps of the past, present, or future. An instructor may also administer values affirmation tasks, which have been shown in some, but not all situations to decrease stereotype threat (Miyake et al., 2010; Lauer et al, 2013). To help women in math courses, instructors may reframe the threat of stereotype as a challenge (Alter, Aronson, Darley, Rodriguez, 2010) and encourage a coping sense of humor in the classroom, which also protects women's math performance from stereotype threat (Ford, Ferguson, Brooks, Hagadone, 2004). All instructors can educate students that intelligence is malleable and should expect greater academic performance, greater engagement, and higher GPA's than those students who feel that intelligence is fixed (Aronson, Fried, & Good, 2001). Over seventeen years ago, Steele called for reducing the interfering pressure of stereotype threat that some students experience, in order to help the overall academic achievement (Steele, 1997) and it is time that our research gave us definitive answers to do so.

Stereotype threat is a complex and difficult area of research to study because one is only able to empirically examine one or two aspects of the environment and individual at a time. These investigations are often in unnatural settings like a psychology lab, despite the fact stereotype threat could be any one of a combination of elements that creates a "threat in the air" (Steele, 1997) for some students. Multiple components of stereotype threat have been discussed, including elements needed to induce stereotype threat, an investigation of how stereotypes are created, possible mechanisms that explain why stereotype threat occurs, and upon whom it is most likely to act. This paper also discusses activation triggers of stereotype threat, possible mechanisms for how it impacts physiology, and how cognition can become distorted as a result of stereotype threat via cognitive dissonance theory. Lastly, possible best practices were discussed for educators who may wish to decrease or eradicate stereotype threat in their classrooms.

It is important to remember that stereotype threat is not felt by all students to the same degree, and assuming achievement gaps exist within one's classroom without proper analysis is unwarranted. In attempting replication of previous work by Miyake et al., (2010), Lauer et al., (2013), investigated the effectiveness of one of the known remedies of stereotype threat, values-affirmation tasks, in order to examine the robustness of the Miyake paper. It was discovered that the women in the science classrooms did not show a gender achievement gap, though the actual sample of the women within some of those science classes, namely physics, was quite small.

What is interesting is that these women were doing just as well as the men in the class, and the endorsement measure of stereotype threat replicated from the Miyake study possibly indicated that there was no threat present. As stated earlier, an instructor should never assume an achievement gap or stereotype threat to be automatic in a STEM classroom, but realize rather, that each individual within that unique environment is capable of equal accomplishments, and should be instructed as such. An area of particular interest for future research in stereotype threat, however, is the field of biology.

Investigations of an achievement gap in undergraduate introductory biology courses have had mixed results (e.g., Willoughby & Metz, 2009; Lauer, 2013), eventhough women routinely underperform on the Medical College Admissions Test (American Association of Medical Colleges, 2012). As more and more calls for introductory biology courses to become more quantitative (e.g., AAAS, 2010), are the women in these courses at risk due to the stereotype that men are better at both math and science than women? Assessing the degree to which women in these courses endorse stereotype threat has thus far, proven difficult. The perception of biology as a “life” science rather than a “physical” science has ramifications for the students enrolled. Currently in biology women and men are at parity with enrollment and graduation rates (NSF, 2011); however, women at times do not do as well as their male counterparts in these biology courses, as evinced by lower normalized learning gains (Willoughby & Metz, 2009). As life science becomes more quantitatively oriented, perhaps an entire sector of women will begin to leave the only area of science where they have had good representation in the past. Since courses like introductory biology are the gateway to upper level courses in botany, ecology, and biochemistry, they eventually will lead to stable, high-paying jobs in STEM fields including the medical sector and biotechnology. Therefore, precise investigation of who is affected by stereotype threat, under what circumstances is it triggered, what classroom atmosphere best ameliorates stereotype threat, and how it can be prevented is of primary concern to prevent a future gush of women out of the biology pipeline, which is so integral to women in STEM.

Currently, women tend to leave the fields of science, technology, engineering, and mathematics in large numbers, which, according to Blickenstaff (2005) is a result of one or many possible reasons (e.g., biological differences between men and women, science pedagogy favoring male students, cultural pressure put on by traditional gender roles, etc.). The small amount of women in science courses like physics shows that those women who persist in college pre-engineering classrooms are somehow special to have continued this far in "the leaky pipeline". By more faculty members being aware of the triggers of stereotype threat and the best practices to avoid it in their classrooms, faculty bias at the undergraduate level may change, showing that women and minorities are more competent than previously thought, which may open doors for women to start the ascent to the top of the professional ladder. Stereotype threat may eventually dissipate as more women persist and succeed in fields of STEM. More women in STEM will be encouraging role models for younger female recruits, alleviating lingering stereotype threat. More women will become domain and socially identified with STEM, as they will move out of the "token" or "minority" positions (Kanter, 1977) in classrooms and the workplace. If women in STEM become the norm, rather than the exception, it will change or eradicate the stereotype that women are not as good at science and math as men. If women are more domain identified, but the stereotype no longer exists, stereotype threat cannot exist, and women will persist.

More research is needed to investigate all variables and contexts that promote the conservation of these women in STEM programs. In addition to answering these questions, more replication studies are needed to lend more credence to the findings in order to strengthen this field of study. As mentioned earlier, the DBER report released by the National Academies of Science calls for more replication studies to fully understand which male or female undergraduate life science students' performance will be affected by stereotype threat and what educators can do to prevent it. At this point, the understanding of the direct mechanisms, who will be affected, who will not be affected, under what circumstances and where is not clearly understood, though this area of research has been intensely studied for nearly 20 years. This area of research is not to the point of knowing which individuals are going to be affected, by what amount, and during what treatment in the laboratory or situation in the classroom, and thus have little to no predictive power. That means the meta-analysis of literature cited above are at best, situational guesses. Until we can definitively say who, what, where, when, why, and how and with a high degree of certainty, this issue is still very much open and in need of more research.

References

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C., & Ruble, D. N. (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology*, *46*(1), 166-171.

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, *40*(3), 401-408.

Aronson, J. & Dee, T. (2012). Stereotype threat in the real world. In M. Inzlicht and T. Schmader (Eds.), Stereotype Threat: Theory, Process, and Application (264-280). New York: Oxford University Press.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113-125.

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim, C. Stangor (Eds.), Prejudice: The target's perspective (83-103). San Diego, CA: Academic Press.

Aronson, J., Lustina, M. J., Good, C., & Keough, K. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology. 35*, 29-46.

Aronson, J., & Salinas, M. F. (1997). Stereotype threat, attributional ambiguity, and Latino underperformance. *Unpublished manuscript, University of Texas*.

Beaton, A., Tougas, F., Rinfret, N., Huard, N., & Delisle, M. N. (2007). Strength in numbers? Women and mathematics. *European Journal of Psychology of Education*, *22*(3), 291-306.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, *12*(3), 225-229.

Blickenstaff, C. J. (2005). Women and science careers: leaky pipeline or gender filter?. *Gender and Education*, *17*(4), 369-386.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307-1310.

Cohen , G.L. Steele, C. M., & Ross, L. (1997). Giving feedback across the racial divide: Overcoming the effects of stereotypes. Unpublished manuscript, Stanford University.

Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. *Personality and Social Psychology Bulletin, 20,* 503-513.

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin. 24*(6), 588-594.

Croizet, J. C., Després, G., Gauzins, M. E., Huguet, P., Leyens, J. P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, *30*(6), 721-731.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*(12), 1615-1628.

Delisle, M. N., Guay, F., Senécal, C., & Larose, S. (2009). Predicting stereotype endorsement and academic motivation in women in science programs: A longitudinal model. *Learning and Individual Differences*, *19*(4), 468-475.

Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford University Press.

Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004). Coping sense of humor reduces effects of stereotype threat on women’s math performance. *Personality and Social Psychology Bulletin*, *30*(5), 643-653.

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C. & Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE-Life Sciences Education*, *6*(2), 132-139.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological bulletin*, *132*(5), 692.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, *28*(5), 659-670.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645-662.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of personality and social psychology*, *85*(3), 541-553.

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213-1216.

Hewlett, S. A., Buck Luce, C., Servon, L., Sherbin, L., Shiller, P., Sosnovich, E., & Suremberg, K. (2008). The Athena Factor: Reversing the Brain Drain in Science, Engineering, and Technology, Boston, MA: Harvard Business School.

Hill, C., Corbett, C., & St. Rose, A. (2010). Why so few? Women in Science, Technology, Engineering, and Mathematics, Washington, DC: American Association of University Women.

Hogg, M. A. (2006). Social identity theory. *Contemporary social psychological theories*, 111-136.

Howard, J., & Hammond, R. (1985). Rumors of inferiority. *New Republic*, *72*(9), 18-23.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, *11*(5), 365-371.

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of experimental psychology. General*, *137*(4), 691.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing Is Half the Battle Teaching Stereotype Threat as a Means of Improving Women's Math Performance. *Psychological Science*, *16*(3), 175-179.

Kanter, R. M. (1977). *Men and Women of the Corporation.* New York: BasicBooks.

Kanter, R. M. (1989). *When giants learn to dance*. New York, NY: Simon & Schuster.

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, *29*(3), 371-381.

Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, *52*(7-8), 489-496.

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80,* 942-958.

Kreutzer, K., & Boudreaux, A. (2012). Preliminary investigation of instructor effects on gender gap in introductory physics. *Physical Review Special Topics-Physics Education Research*, *8*(1), 010120.

Langer, E. J., Bashner, R. S., & Chanowitz, B. (1985). Decreasing prejudice by increasing discrimination. Journal of Personality and Social Psychology, 49, 113–120.

Latane, B. (1981). The psychology of social impact. *American psychologist*, *36*(4), 343-356.

Lauer, S., Momsen, J., Offerdahl, E., Kryjevskaia, M., Christensen, W., & Montplaisir, L. (2013). Stereotyped: Investigating Gender in Introductory Science Courses. *CBE-Life Sciences Education*, *12*(1), 30-38.

Lawrence, J. S., Crocker, J., & Dweck, C. S. (2005). Stereotypes negatively influence the meaning students give to academic settings. In G. Downey, J. S. Eccles, and C. M. Chatman (Eds.), *Navigating the future: Social Identity, Coping, and Life Tasks* (23-44). New York: Russell Sage Foundation.

Lazarus, R. S. (1999). Stress and emotion: A new synthesis. New York: Springer Publishing Company.

Lemkau, J. P. (1979). Personality and background characteristics of women in male-dominated occupations: A review. *Psychology of Women Quarterly*, *4*(2), 221-240.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, *39*(5), 821.

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology*, *96*(6), 1089-1103.

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, *38*(4), 405-413.

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183-1193.

McIntyre, R. B., Lord, C. G., Gresky, D. M., Ten Eyck, L. L., Frye, G. J., & Bond Jr, C. F. (2005). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. *Current Research in Social Psychology*, *10*(9), 116-36.

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*(1), 83-90.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science*, *330*(6008), 1234-1237.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the USA*, *109*, 16474-16479.

Myers, D. G. (2004). *Psychology: Seventh edition in modules* (7^{th} ed.). New York: Worth Publishers.

National Science Foundation (2011). Women, Minorities, and Persons with Disabilities in Science and Engineering: 2011 (No. Special Report NSF 11-309), Arlington, VA.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annu. Rev. Econ.*, *3*(1), 601-630.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of personality and social psychology*, *83*(1), 44-59.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., & Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593-10597.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*(6), 782-789.

Olson, S., & Riordan, D. G. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Math. Report to the President. *Executive Office of the President*.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of personality and social psychology*, *76*, 114-128.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194-201.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of personality and social psychology*, *85*(3), 440-452.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological review*, *115*(2), 336.

Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology*, *24*(6), 490-504.

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research, 20*(4), 418-432.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4-28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist. 52*(6), 613-629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-797.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in experimental social psychology*, *34*, 379-440.

Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology.* Cambridge, UK: Cambridge University Press.

Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behavior. *S. Worchel & WG Austin (red.), The Psychology of Intergroup Relations*, 7-24.

Walton, G. M., & Spencer, S. J. (2009). Latent ability grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*(9), 1132-1139.

Watson, B. (2013, March 16). The High Cost of Higher Education Explained in One Simple Graphic. Retrieved from http://www.dailyfinance.com/on/college-costs-tuition-rising-student-debt-infographic/

Willoughby, S. D., & Metz, A. (2009). Exploring gender differences with different gain calculations in astronomy and biology. *American Journal of Physics*, *77*, 651.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*, *18*(5), 459-482.