The Report Committee for Alyssa Lynne Reinhart
Certifies that this is the approved version of the following report:

Revisiting Rydell, McConnell, & Beilock (2009): A case for the inclusion of non-targets of stereotype threat

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: ______________________________
Gary D. Borich

____
Keisha Bentley-Edwards

Revisiting Rydell, McConnell, & Beilock (2009): A case for the inclusion of non-targets of stereotype threat

by

Alyssa Lynne Reinhart, B.S. Math.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2014

Acknowledgements

I would like to thank my supervisor, Dr. Gary Borich, and reader, Dr. Keisha Bentley-Edwards for their guidance and support. I would also like to thank Dr. Erika Patall for her feedback on an earlier version of this report.

Abstract

Revisiting Rydell, McConnell, & Beilock (2009): A case for the inclusion of non-targets of stereotype threat

Alyssa Lynne Reinhart, M.A.

The University of Texas at Austin, 2014

Supervisor: Gary D. Borich

This study sought to examine the role of multiple identities as a possible protective factor against stereotype threat for females taking a difficult math test. Specifically, it sought to replicate the findings of Rydell, McConnell, and Beilock (2009), who found that making a positively stereotyped identity salient (college) at the same time a negatively stereotyped identity (female) was salient, buffered the effects of stereotype threat. This study also attempted to evaluate the validity of a common experimental stereotype threat manipulation, which is to make explicit statements about performance which remind test subjects of existing stereotypes.

Using a quantitative experimental design, and replicating the methodology used in the 2009 study, math-identified college students were randomly assigned to take a difficult math test under circumstances which varied salient identities. For the experimental conditions, an explicit statement was made about prior performance by either females or females and college students. For math-identified females, the statement about female performance was believed to invoke a negative stereotype about math

ability and thus stereotype threat. However, when the statement was about both their gender as well as their college identity (thought to be positively stereotyped), this would cause the females to suppress their gender identity in order to maintain positive self-esteem and thus would be protected from stereotype threat effects. It was also predicted that non-targets of threat (males) would not be affected by the manipulations, as according to the theory of stereotype threat, a stereotype has to be self-relevant to become a threat.

Results failed to replicate the findings of the previous study. While not significant, females actually trended towards better performance when reminded of the negative stereotype about females, as compared to a control group. More importantly, this type of manipulation was shown to significantly affect non-targets of threat, which is a violation of stereotype threat theory. When reminded of the negative stereotype about females, males performed significantly worse than a control group. This evidence supports the idea that making explicit statements about ability is an invalid method of invoking stereotype threat in an experimental setting.

TABLE OF CONTENTS

Introduction

"For whatever reason, I didn't succumb to the stereotype that science wasn't for girls. I got encouragement from my parents. I never ran into a teacher or a counselor who told me that science was for boys. A lot of my friends did." – Sally Ride

What does it take for a female to pursue a career in the math and sciences? Why do some females succeed in the face of negative stereotypes about their abilities in these fields? According to the Department of Education, despite gains in overall educational enrollment, women make up less than 25% of participants in science, technology, engineering, and math (STEM) fields (United States Department of Education, 2012). Is this because these fields are seen as areas where men have greater ability? It has actually been shown that when reminded of their gender, women indicated they were more interested in arts over math (consistent with the stereotype, and in contrast to a control group; Steele & Ambady, 2006). This change in behavior due to the existence of a negative stereotype about a part of one's identity has been attributed to a phenomenon known as stereotype threat (Steele & Aronson, 1995; Wheeler, DeMarre, & Petty, 2004; Steele & Ambady, 2006). Stereotype threat occurs when one is trying to not confirm a negative stereotype about an aspect of their identity, which can cause decreased performance and possibly disidentification with a domain entirely (Steele, 1997; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Spencer, Steele, & Quinn, 1999).

Recent research has suggested that simultaneously activating a positively stereotyped aspect of one's identity, while another negatively stereotyped identity is salient, could buffer from the effects of stereotype threat (Shih, Pittinsky, & Ambady, 1999; Rydell, McConnell, & Beilock, 2009). However, there is still disagreement in the stereotype threat literature on how positively stereotyped identities affect performance under threat (Cheryan & Bodenhausen, 2000; Shih, Pittinsky, & Ambady, 1999; Rydell, McConnell, & Beilock, 2009). Furthermore, there is no consensus on how to manipulate threat in experimental settings.

Research has suggested that repeatedly experiencing this threat can over time lead to disidentification with a domain altogether (Steele & Aronson, 1995; Steele, 1997;; Spencer, Steele, & Quinn, 1999; Cokley, 2002). These findings lead one to question whether stereotype threat may be a contributing factor to the gender gap in STEM career pursuit. If we can identity protective factors which females could invoke under stereotype threat, perhaps this would decrease the likelihood that they leave science and math fields. For example, could something as simple as reminding a student of a positively stereotyped aspect of themselves, such as belonging to a competitive university, help her relax in the moment and perhaps even perform better?

More research is needed in order to not only further understand what these protective factors are and how they work, but to also establish a consistency in the literature regarding how stereotype threat is studied. Once we can ensure we are indeed studying threat, and not another phenomenon, we could move towards designing interventions. This study sought to examine how invoking multiple identities under threat would affect performance on a math test by replicating the findings of Rydell, McConnell, and Beilock (2009), while at the same time testing the validity of the threat manipulation by including non-targets of threat. It also proposes a program to study and evaluate this phenomenon and possible intervention in the classroom.

Background

STEREOTYPE THREAT

Despite numerous educational initiatives, there continues to be a gap between males and females in the pursuit of a career in science, technology, engineering, and mathematics (STEM; Spencer, Steele, & Quinn, 1999; Steel & Ambady, 2006). A possible explanation for this is that when certain negative stereotypes are primed and they are personally relevant, it leads to a change in behavior (Wheeler, DeMarre, & Petty, 2004; Steele & Ambady, 2006). This change can even decrease the performance on domain specific tasks, and after a time it can even lead to disidentification with the domain altogether (Steele & Aronson, 1995; Steele, 1997; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Spencer, Steele, & Quinn, 1999). This phenomenon, in which awareness of a negative stereotype about a domain which one identifies with (females in math, Blacks in academics, etc.,), and the subsequent decrease in performance is known as *stereotype threat*.

Stereotype threat is defined as "being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele & Aronson, 1995). This self-evaluative threat was first identified by Claude Steele and Joshua Aronson (1995) in a study examining the performance of African American students. The authors suggested that something, perhaps chronic exposure to negative stereotypes about the academic ability of Blacks, was behind a trend of lower academic performance as well as standardized tests over-predicting subsequent achievement (relative to Whites with the same academic preparation). Through a series of four experiments, the authors demonstrated that performance on an identical task varied between groups when the purpose of the task was framed differently – the first group was told the test was just an exercise, while the second was told it was a test of intelligence. Steele and Aronson found that the group who was told that the test was diagnostic of their intelligence performed significantly worse.

Since the publication of this seminal piece in 1995, the field of research on stereotype threat has exploded to look at everything from athletic performance (Stone, Lynch, Sjomerling, & Darley, 1999; Stone & McWhinnie, 2008) to memory performance in the elderly (Levy, 1996), investigating not only the underlying causes of the effect but also ways to counteract it. It has been proposed that stereotype threat may not only contribute to gender differences on test performance, but may also explain gender differences in career choices. For example, when reminded of their gender, females are less likely to express an interest in math academic domains over the Arts (Steele & Ambady, 2006). Steele suggested that chronic exposure to a negative stereotype could eventually cause disidentification with the domain (1997). This could explain why many females choose to not pursue a career in the STEM fields, because of the persistent stereotype that they lack the ability. Some females however, such as Sally Ride, have managed to overcome this stereotype. Recent research has suggested that invoking a positive stereotype about another aspect of one's identity could counteract threat, but the results are mixed (Cheryan & Bodenhausen, 2000; Shih, Pittinsky, & Ambady, 1999; Rydell, McConnell, & Beilock, 2009).

STEREOTYPE THREAT AND IDENTITY

According to social identity theory (Tajfel & Turner, 1986; Rydell, McConnell, & Beilock, 2009), humans want to maintain a positive image of themselves. If confronted with a negative stereotype about a part of their identity, they may leave that group (suppress that identity) and increase their identity salience with a more positively viewed group. For example, an Asian American female who identifies with the math domain is aware of both the negative stereotype that females are not strong in math relative to men, as well as the positive stereotype that Asians are the best at math compared to other ethnicities. Because she wants to think positively of herself, she may suppress her female identity and make her Asian identity salient. This logic has not been consistently supported in the literature however.

Positive Identity Saliency

Using a population of Asian American women, Shih, Pittinsky, and Ambady (1999) conducted two experiments which studied implicit stereotype threat effects on mathematics test performance when different identities were made salient. To make the identities and stereotypes salient, the subjects answered general questions about either their gender or their ethnicity (or neither) before completing the task. When gender was made salient, the subjects performed worse on a quantitative exam relative to a control group. But when their Asian identity, and therefore the stereotype that Asians are good at math, was made salient, the subjects performed better than the control group.

In 2000, Cheryan and Bodenhausen built off of this work but used manipulations which they felt were more appropriate for making an identity salient. Before completing a math test, the female Asian American subjects answered questions from a modified version of the Collective Self-Esteem Scale (Luthanen & Crocker, 1992) which had them indicate the degree to which they agreed with statements such as, “I am a worthy member of the gender that I belong to,” and “Overall, my race is considered good by others.” Unlike the previous study (Shih, Pittinsky, & Ambady, 1999), when ethnic identity was made more salient, these subjects actually performed worse than the control group; while the gender condition did not vary significantly from the control group. Furthermore, participants in the ethnicity condition were more likely than the control group to report difficulty concentrating. This trouble focusing partially mediated the effect of threat on performance for the ethnicity condition.

Simultaneous Saliency

Rydell, McConnell, & Beilock (2009) extended this line of research by examining how explicitly activating multiple identities, specifically one positively and one negatively stereotyped identity, at the same time could change how stereotype threat affects performance. Before taking a difficult math test, a group of female college students was primed with either a negative statement regarding performance on the test by females, both the negative statement about females as well as a positive statement about performance by college students, or nothing. The authors found that the participants who only saw the negative statement performed more poorly on the math test than the control group. However, participants in the condition which saw both the negative and positive statement performed better than those in the negative condition and also showed lower gender identity accessibility. Using social identity theory, the authors argued that the females had suppressed their negatively stereotyped identity (female) and made the positively stereotyped identity (college) more salient, which in turn led to a higher performance on the math test.

These three studies triggered identity salience in different ways, which brings into question, what is a valid manipulation of threat in experimental studies? One of the major criticisms of stereotype threat research is that many researchers may in fact be studying an entirely different phenomenon, known as stereotype priming.

STEREOTYPE THREAT OR STEREOTYPE PRIMING

Stereotype priming occurs when a person takes on the characteristics of a primed stereotype, but unlike stereotype threat, this stereotyped identity does not have to be self-relevant (Steele & Aronson, 1995; Marx, 2011). For example, when college students were primed with traits of a university professor, they performed better on a test of general knowledge than students primed with traits of "hooligans" (Dijksterhuis & van Knippenberg, 1998). Another study found that after implicitly primed with stereotyped traits about the elderly, college students walked down a hallway at a significantly slower pace than a control group (Bargh et al., 2001; Marx, 2011).

Could stereotype priming help explain the difference in results between the three studies on the effects of positive stereotypes? Furthermore, how do we ensure that we are not priming? As per Claude and Aronson's original definition of threat, the stereotype must be self-relevant. If it is not self-relevant, then there should be no effect. Therefore, the best way to ensure that we are studying threat and not priming is to ensure the subjects identify with the stereotyped domain as well as include a control group of participants for which the threat manipulation is not relevant. Not only do the three previously mentioned studies fail to determine how closely the participants identified with a math domain, they also failed to include a control group of non-targets of threat.

Current Study

To further this line of research, the current study sought to replicate the work done by Rydell, McConnell, and Beilock (2009), while also including non-targets of threat, namely White males. Measures of math and academic (college) identity were included as well to ensure the manipulations were relevant. This study specifically sought to answer the following questions:

Research Question 1: *When positive and negative stereotypes about domain ability are concurrently salient, how is performance of threat targets affected?*

Hypothesis 1: It was hypothesized that the data would confirm the previous findings (Rydell, McConnell, & Beilock, 2009), in that math-identified females would perform better when both a positive and negative stereotype were present.

Research Question 2: Is there evidence to ensure that this experimental manipulation of threat is different from stereotype priming?

Hypothesis 2: It was predicted that the manipulations would have no effect on the males.

METHOD

Participants

One hundred thirty-seven students, 63 males and 74 females, from a large university in the southern part of the United States participated in the study. The sample consisted of 35.3% White, 29.3% Hispanic, and 27.1% Asian American, with the remaining spread about equally between African American (2.3%), multiracial (3.8%), or other (2%). The students came predominately from the colleges of liberal arts, natural science, and business, and were mostly sophomores, juniors, and seniors. Upon approval of the Institutional Review Board, the participants were recruited from a departmental research subject pool and received credit in one of several core educational psychology courses for their participation.

Outcome Measures

Participants completed twelve quantitative reasoning questions drawn from two sample GRE tests, obtained from the Educational Testing Service (ETS) website (2009a, 2009b). These multiple-choice mathematic questions were chosen because they had only one answer and assessed each participant's range of abilities in arithmetic, algebra, geometry, and data analysis. A score was calculated for each participant by determining the percent of questions that they answered correctly.

Identity. In order to ensure that the manipulations were self-relevant, the subjects completed measures on both college identity and math identity. Was & Isaacson's (2008) Academic Identity Measure (AIM) was used to determine how closely the participants identified with a school-going culture. The AIM is based on an idea that like other identities, there is a process of development. Its 40 Likert-scale items measure what stage in the development of an academic identity one is in: Moratorium (someone who is in transition), Foreclosed (someone whose reasons for pursuing academics are mostly due to the expectations of others, namely family and friends), Diffuse (someone who does not identify at all with academics), or Achievement (someone who identifies highly with academics purely for their own interests). This measure has good internal consistency with alphas at .77 for foreclosed, .76 for achievement, .76 for diffusion, and .85 for moratorium. Scores for each of these subscales were calculated to determine which of the four styles, the participants ranked highest. Mathematics domain identity (Walton, 2008) was measured by how closely the participants aligned with a mathematics identity. It consisted of 14 Likert-scale items which measured participants in two domains, beliefs in their math abilities ("It is possible for me to get good grades in math.") and their self-perception as related to math ("I think like a mathematician."). In prior research, this measure has reliability of .83 for math abilities and .76 for math self-perception. A mean score was calculated for each participant across the fourteen items to give them a general Math Identity score with a range of (14-56).

Experimental Procedures

The study was conducted in a university computer lab, where participants completed a series of measures as well as GRE questions via an online survey using the Qualtrics survey system. The experiment was facilitated by a female researcher. Subjects first completed a series of background measures to determine their gender, ethnicity, and general demographics.

Stereotype Threat. Based on prior research (Rydell, McConnell, & Beilock, 2009), participants were then randomly assigned to one of three stereotype threat conditions: 1) neutral/control, 2) negative stereotype, or 3) positive and negative stereotype). The control condition stated that the students would be completing some practice GRE-like questions in an effort to ensure that the questions did not have any errors. The negative stereotype threat indicated that students were about to complete GRE questions that historically females performed worse on, which was consistent with the idea that females had lower math ability. The positive and negative stereotype threat condition repeated the statements from the negative threat condition, but that college students performed quite well on them.

Outcome. The participants completed twelve quantitative GRE questions, which gave immediate feedback after each question, including the correct response as well as their current score. After being exposed to their assigned stereotype threat condition, the participants were asked to estimate how many points they believed they would have at the end of the test (due to technical difficulties with the online survey during one of the sessions, thirteen participants were unable to complete this section). After completing the GRE section, participants completed measures of academic identity and math domain identity. These measures were placed at the end of the manipulation so that it more closely replicated the manipulation in the Rydell, McConnell, & Beilock (2009) study.

This was done because asking about identity before the test could be seen as a manipulation of identity salience.

Analytical Strategy. Descriptive statistics by group were calculated for each measure. Then an ANOVA, gender by condition, was conducted to look for between-group differences. The following section details the findings of this analysis. Only participants who indicated a math domain identity were included in the analysis.

Results

Experimental Groups. The number of participants by experimental condition and reward structure is shown in Table 1.

Table 1. *Participants by Group*

	Control	**Gender**	**Gender & College**
Female	Gains: 12	Gains: 14	Gains: 14
	Losses: 15	Losses: 8	Losses: 11
Male	Gains: 7	Gains: 12	Gains: 10
	Losses: 11	Losses: 11	Losses: 12

Baseline Data. There were no significant differences between grade levels or colleges, so these were analyzed as a cohesive group. As seen in Table 2, there were no significant differences between genders on any of the measures. However, males were fairly accurate in predicting their performance, while females expected to perform worse than they actually did.

Table 2. *Comparison of Scores on Math Domain Identity, Academic Identity, and Expectation/Performance by Gender*

Measure	Gender	Mean	SD	(Min, Max)
Math Domain Identity	Male	47.22	9.75	(22.00, 70.00)
	Female	42.53	9.14	(8.33, 91.67)
Academic Identity – Achieve	Male	36.13	5.79	(20.00, 46.00)
	Female	37.69	5.47	(23.00, 48.00)
Academic Identity – Diffusion	Male	25.40	6.19	(11.00, 41.00)
	Female	22.08	5.09	(12.00, 32.00)
Academic Identity – Foreclosure	Male	30.27	6.62	(14.00, 42.00)
	Female	30.43	6.17	(19.00, 46.00)
Academic Identity – Moratorium	Male	30.65	5.59	(14.00, 40.00)
	Female	30.70	6.17	(10.00, 44.00)
Difference in Expectation vs. Actual Performance on Quantitative	Male	0.12	5.72	(-17.00, 16.00)
	Female	-3.85	6.73	(-21.00, 15.00)

Hypothesis Testing. It was predicted that when females were exposed to only the negative stereotype, their performance would decrease relative to the control group. It was also predicted that when females were exposed to both the positive and negative stereotypes, they would suppress the negative identity, and thus would perform as well as the control group and better than the negative stereotype group. As shown in Table 3, an ANOVA determined that gender and the interaction between gender and experimental condition were statistically significant at a $p=.05$ level, while experimental condition was marginally significantly ($p=.10$) related to scores on the math test. Furthermore, post-hoc tests found that math-identified females performed significantly better when they were exposed to the negative statement about female ability in math as compared to the control group ($p<.05$). While females performed second best in the condition where both the positive and negative stereotype was activated, this was not significantly different from either the control or the negative stereotype group.

Table 3. *ANOVA Table*

Source	**df**	**F**	**η**	**p**
Condition	2	2.37	0.20	0.10
Gender	1	4.06	0.19	0.05
Condition*Gender	2	3.71	0.25	0.03

Math-identified males on the other hand performed best in the control group, and this was significantly better than the dual stereotype condition ($p<.01$). They performed second best under the negative stereotype condition but not significantly better than the control. These results are shown in Table 4.

Table 4. *Condition*Gender Comparisons*

Condition	**Gender**	**Mean**	**Std. Error**	**95% Confidence Interval**
Control	Male	80.35	4.86	(70.73, 89.98)
	Female	59.26	4.29	(50.76, 67.77)
Negative Stereotype	Male	75.34	4.41	(66.61 84.07)
	Female	71.36	4.43	(62.59, 80.13)
Positive & Negative	Male	62.73	4.39	(54.03, 71.44)
Stereotype	Female	65.05	4.37	(56.39, 73.70)

Discussion

The purpose of this study was to replicate the experiment conducted by Rydell, McConnell, and Beilock (2009) but to include non-targets of threat, namely males to ensure that the manipulation they used was relevant to the study of stereotype threat. Specifically, it looked at how explicitly invoking multiple identities under stereotype threat would affect performance on a difficult math test. It was found that females performed best when confronted with the explicit negative stereotype. Additionally, males performed best under the control condition, but worse when the stereotypes were introduced.

In Rydell, McConnell, and Beilock's (2009) article, they found that when explicitly confronted with both a negative stereotype about math performance – that college students performed well on tests of quantitative ability but females did not, females performed better on a math test as compared to when they were only confronted with the negative stereotype. This attempt at replication not only failed to repeat their results, but also showed that the manipulation affected a non-target group. They argued that based on social identity theory, we want to have a positive opinion of ourselves and when confronted with two aspects of our identity, one associated with a positive stereotype and one with a negative stereotype, that we will focus on the positively stereotyped identity and suppress the negatively stereotyped one.

Furthermore, a cognitive imbalance occurs when a stereotype about part of your identity conflicts with what you know about yourself. Therefore, if you are a female who identifies strongly with math but are confronted with a negative stereotype about female math performance on standardized tests, this does not fit logically with what you know of yourself, which leads to self-doubt, a heightened vigilance for not confirming the negative stereotype, and thus a decrease in performance (Schmader et al., 2008, Johns & Schmader, 2010; Schmader & Beilock, 2012). In the current study however, females did not perform significantly better in the condition where both positive and negative stereotypes were presented. In comparison to the control condition, females who only saw the negative stereotype performed significantly better; this performance was also better than those who saw both positive and negative stereotypes, but that difference was not significant. One potential explanation for these contrasting results is that the participants had ample time to complete the quantitative items. If the threat induced a state of hyper vigilance, especially to that of preventing errors, the time factor could have compensated for the fact that this concentration on accuracy slowed the females down. A future study should vary the time allowed to complete the math questions in order to rule this out. However, this could also be interpreted as a non-valid manipulation of threat. To further that claim, it is important to consider the effect on non-targets of threat.

Surprisingly, males performed the poorest in the condition where they were presented with both the positive and negative stereotypes. Could this outcome also be stereotype priming? In this condition males performed significantly worse than both the control condition as well as the condition which only saw the negative stereotype about females and math. If a decrease in performance is due to a cognitive imbalance, which in turn is caused by a mismatch in self-perception and a stereotype related to an aspect of one's identity, it fails to really explain this difference. When a male college student who identifies with a math domain is confronted with a negative stereotype about female math performance, according to social identity theory, it should not affect him, as being female is not a part of his identity. However, the results from this study do not support this logic. Did the males interpret the stereotypes as pressure on themselves to perform well? Previous work on a related phenomenon known as stereotype lift has shown that performance of non-targets can actually increase when they are aware of a negative stereotype about a group which they are not a member of (Walton & Cohen, 2003). For example, it has been shown that when taking a math test which is known to produce gender differences, females tend to perform worse than a control group while males will perform better (Shih, Pittinsky, & Ho, 2013). Perhaps this actually occurred in the control groups for this experiment, as this condition had both the lowest scores for females and highest scores for males. However, it fails to fit the other conditions in this experiment which suggests something else, such as stereotype priming was occurring. More research that looks into what happens within genders under stereotype threat is needed.

Conclusion and Future Directions

This study failed to replicate the findings of Rydell, McConnell, & Beilock (2009), and in fact had almost reverse effects. However, through the use of non-targets of threat, in this case males, it supports the idea that explicit manipulation of stereotype threat may be invalid. The fact that males were adversely affected by the negative stereotypes about females is evidence of stereotype priming effects, rather than stereotype threat effects. More research is needed to not only determine if multiple identities are truly a protective factor in the face of stereotype threat (see the proposed program evaluation in Appendix C), but to also further refine how we study threat.

APPENDIX A: MEASURES

Academic Identity Measure (Was & Isaacson, 2008)

1. Good grades have always been important for me because I like to make my parents proud.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

2. Sometimes I think the reason I'm in college is I have nothing better to do.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

3. I'm not sure what occupation I want after college and I'm not really concerned about it yet.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

4. A college education is a high priority for me and I'm willing to make sacrifices.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

5. I've considered a number of college majors and have decided which one is best for me.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

6. I always knew my college major mainly from the guidance I received from my family.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

7. I want a college education but sometimes I'm not sure I can make the commitment.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

8. I don't worry about grades very often and rarely set academic goals for myself.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

9. How I do in school is important to me because others are counting on me to do well.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

10. I've never decided on my own about college. I just did what friends and family expected of me.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

11. My priorities for school come from my early experiences. I usually just accept what is expected of me.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

12. My view of grades and studying fluctuates; sometimes I am conscientious, other times I am lazy.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

13. If I had to pay for my own education, I probably wouldn't be in school even if I had the money.

(1)	(2)	(3)	(4)	(5)

Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

14. Sometimes I feel responsible for my learning but other times I feel it is out of my hands.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

15. In class, my mind often wanders and I often wish I were someplace else.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

16. An important reason I chose to go to college was my family wanted me to go.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

17. If a class is important, I can concentrate even if the teacher or topic is boring.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

18. I feel comfortable being responsible for my education and learning.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

19. Of all of the reasons to be in college, one of my most important reasons is social and friendships.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

20. I feel I have to attend every college class; otherwise my parents would be upset.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

21. Some days I am enthusiastic about learning, but other days I don't really care.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

22. I try to write down everything the professors say but I seldom think about applications.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

23. If a class is very difficult, I will usually give up and blow it off.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

24. My priorities in school are in transition. Some days I am serious, other days I have other priorities.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

25. When I do poorly on a test, I think of what I did wrong and try to solve the problem.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

26. I don't have clear priorities for school and life. I usually just go with the flow.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

27. I want to complete my school work but I often look back and realize I didn't set aside the time.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

28. I find most class topics at least somewhat interesting – I'm seldom bored in class.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

29. If a class is very difficult, I buckle down and study more so I don't disappoint other people.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

30. Although I have many priorities, learning in school is always one of my most important goals.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

31. Sometimes I feel confident I know what I want from my education but other days I'm not so sure.

(1)	(2)	(3)	(4)	(5)
Strongly	Disagree	Neither Agree	Agree	Strongly Agree

Disagree Nor Disagree

32. I know why I am in college and have clear goals that I want to achieve.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

33. When I do poorly on a test, I get upset and worry what friends and family might think of me.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

34. Sometimes I get upset when I do poorly on a test and other times I just let it slide.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

35. Finding time to study often takes a back seat to social and recreational activities.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

36. When a course is demanding, my first reaction is to work harder, but sometimes I give up.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

37. Sometimes I am interested in what is being discussed in class but other days I am bored.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

38. When school is challenging, I find a way to learn even if I have to try new ways to study.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

39. Most of the material I am asked to learn in my classes is boring.

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

40. Finding time to study may be difficult so I set aside time to complete my school work.

(1)	(2)	(3)	(4)	(5)
Strongly	Disagree	Neither Agree	Agree	Strongly Agree

Disagree Nor Disagree

Scoring:

Diffusion (lack of exploration or commitment): 2, 3, 8, 13, 15, 19, 23, 26, 35, 39

Achievement (commitment to a set academic values): 4, 5, 17, 18, 25, 28, 30, 32, 38, 40

Foreclosure (academic identity depends on significant others): 1, 6, 9, 10, 11, 16, 20, 22, 29, 33

Moratorium (time of academic indecision): 7, 12, 14, 21, 24, 27, 31, 34, 36, 37

Mathematics Domain Identity (Walton, 2008)

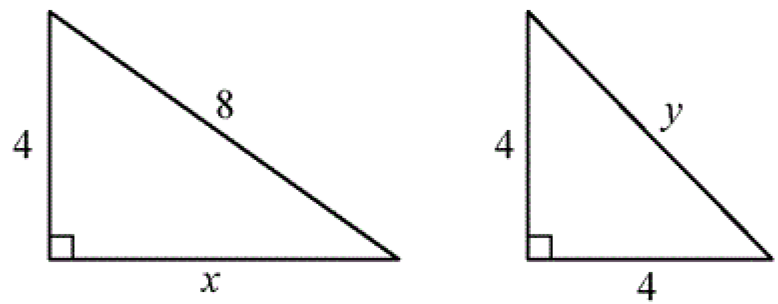
Students have different thoughts and feelings about their math classes. Please read each statement carefully and use the scales to rate your opinions about math and your math class.

For me it is possible to…	Not Possible	Not Very Possible	Somewhat Possible	Possible	Very Possible
Think like a mathematician.					
Be a math helper or tutor.					
Not be able to do the math required for my job when I graduate.					
Get a good grade in a math course.					
Earn poor grades in math courses.					
Be one of the top students in a math class.					
Be afraid to take more math classes.					
Help my friends get good grades in a math class.					
Use my math skills to solve					

problems outside of school.					
Fail a math class.					
Major in math.					
Do poorly on my next math test.					
Become a math teacher.					
Get a job that requires math skills.					

Graduate Record Examination (GRE) Quantitative Questions

Please answer the following questions to the best of your ability. You may use scratch paper if you would like. You will have 40 minutes to complete this section.



1)

Quantity A: x

Quantity B: y

From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

2) It is given that $(x - 2y)(x + 2y) = 4$

Quantity A: $x^2 - 4y^2$

Quantity B: 8

From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

3) A certain recipe requires 3/2 cups of sugar and makes 2 dozen cookies.

Quantity A: The amount of sugar required for the same recipe to make 30 cookies

Quantity B: 2 cups

From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

4) A power station is located on the boundary of a square region that measures 10 miles on each side. Three substations are located inside the square region.

Quantity A: The sum of the distances from the power station to each of the substations

Quantity B: 30 miles

From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

5) It is given that 6 is less than x, which is less than 7, and y = 8

Quantity A: x/y

Quantity B: 0.85

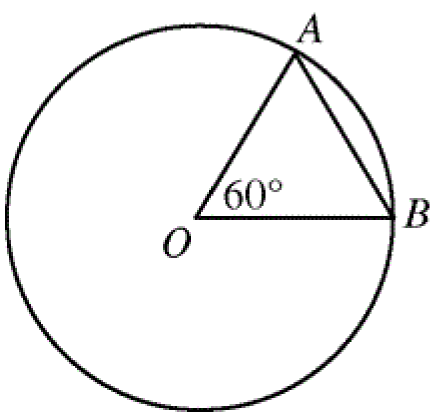
From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.



6)

It is given that O is the center of the circle and the perimeter of triangle B O A is 6.

Quantity A: The circumference of the circle

Quantity B: 12

From the answer choices given, select and indicate the one that describes the relationship between quantity A and quantity B.

A. Quantity A is greater.

B. Quantity B is greater.

C. The two quantities are equal.

D. The relationship cannot be determined from the information given.

7) This question has five answer choices, labeled A through E. Select the best one of the answer choices given.   The system of equations

$7x + 3y = 12$, and

$3x + 7y = 6$ is given.

If x and y satisfy the system of equations given, what is the value of $x-y$?

A. 2/3

B. 3/2

C. 1

D. 4

E. 6

8) This question has five answer choices, labeled A through E. Select the best one of the answer choices given. If $(5^{5x})(25) = 5^n$, where n and x are integers, what is the value of n in terms of x.

A. $5x + 1$

B. $5x + 2$

C. $5x + 5$

D. $10x$

E. $10x + 2$

9) This question has five answer choices, labeled A through E. Select the best one of the answer choices given. In the sunshine, an upright pole 12 feet tall is casting a shadow 8 feet long. At the same time, a nearby upright pole is casting a shadow 10 feet long. If the lengths of the shadows are proportional to the heights of the poles, what is the height, in feet, of the taller pole?

A. 10

B. 12

C. 14

D. 15

E. 18

10) This question has five answer choices, labeled A through E. Select the best one of the answer choices given. If k is the smallest prime number greater than 21 and b is the largest prime number less than 16, then kb =

A. 299

B. 323

C. 330

D. 345

E. 351

11) This question has five answer choices, labeled A through E. Select the best one of the answer choices given.

List R: 28, 23, 30, 25, 27

List S: 22, 19, 15, 17, 20 The median of the numbers in list R is how much greater than the median of the numbers in list S?

A. 8

B. 10

C. 12

D. 13

E. 15

12) This question has five answer choices, labeled A through E. Select the best one of the answer choices given. Each month, a certain manufacturing company's total expenses are equal to a fixed monthly expense plus a variable expense that is directly proportional to the number of units produced by the company during that month. If the company's total expenses for a month in which it produces 20,000 units are $570,000, and the total expenses for a month in which it produces 25,000 units are $705,000, what is the company's fixed monthly expense?

A. $27,000

B. $30,000

C. $67,500

D. $109,800

E. $135,000

APPENDIX B: TABLES

Table 1. *Participants by Group*

	Control	**Gender**	**Gender & College**
Female	Gains: 12	Gains: 14	Gains: 14
	Losses: 15	Losses: 8	Losses: 11
Male	Gains: 7	Gains: 12	Gains: 10
	Losses: 11	Losses: 11	Losses: 12

Table 2. *Comparison of Scores on Math Domain Identity, Academic Identity, and Expectation/Performance by Gender*

Measure	Gender	Mean	SD	(Min, Max)
Math Domain Identity	Male	47.22	9.75	(22.00, 70.00)
	Female	42.53	9.14	(8.33, 91.67)
Academic Identity – Achieve	Male	36.13	5.79	(20.00, 46.00)
	Female	37.69	5.47	(23.00, 48.00)
Academic Identity – Diffusion	Male	25.40	6.19	(11.00, 41.00)
	Female	22.08	5.09	(12.00, 32.00)
Academic Identity – Foreclosure	Male	30.27	6.62	(14.00, 42.00)
	Female	30.43	6.17	(19.00, 46.00)
Academic Identity – Moratorium	Male	30.65	5.59	(14.00, 40.00)
	Female	30.70	6.17	(10.00, 44.00)
Difference in Expectation vs. Actual Performance on Quantitative	Male	0.12	5.72	(-17.00, 16.00)
	Female	-3.85	6.73	(-21.00, 15.00)

Table 3. *ANOVA Table*

Source	**df**	**F**	**η**	**p**
Condition	2	2.37	0.20	0.10
Gender	1	4.06	0.19	0.05
Condition*Gender	2	3.71	0.25	0.03

Table 4. *Condition*Gender Comparisons*

Condition	Gender	Mean	Std. Error	95% Confidence Interval
Control	Male	80.35	4.86	(70.73, 89.98)
	Female	59.26	4.29	(50.76, 67.77)
Negative Stereotype	Male	75.34	4.41	(66.61 84.07)
	Female	71.36	4.43	(62.59, 80.13)
Positive & Negative Stereotype	Male	62.73	4.39	(54.03, 71.44)
	Female	65.05	4.37	(56.39, 73.70)

APPENDIX C: PROGRAM EVALUATION PROPOSAL

Program Evaluation Proposal

Background

Female retention and graduation rates in science, technology, engineering, and mathematics (STEM) degree programs continue to be a major problem (United States Department of Education, 2012). Research suggests that constant exposure to stereotype threat could be one reason for this trend; repeated exposure to a negative stereotype about a part of one's identity can lead to disidentification with that domain (Steele & Aronson, 1995; Steele, 1997; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Spencer, Steele, & Quinn, 1999). A study conducted in 2008 by Good, Aronson, and Harder found that by indicating a test was gender neutral (or gender-fair), or that it did not show gender differences in performance, the researchers were able to nullify the effects of stereotype threat. Could elimination of stereotype threat on course examinations help address the female retention problem in STEM fields?

Statement of Purpose

This program evaluation seeks to answer the following questions: Do female students perform better on exams when told the tests are gender-fair? Do female students who continue to take these "gender-fair" tests persist longer in their STEM majors and have higher graduation rates? This evaluation consists of one first order outcome, two second order outcomes, and one third order outcome. The first order outcome is directly related to exam and course grades and will be compared to previous, historical data, as well as other sections of the same courses that will not be given the "gender-fair" tests (control group). The second order outcomes will be the females' STEM identity beliefs after the course, compared to scores on initial STEM identity measurements administered at the beginning of the course, as well as data on whether or not they stayed in their major, compared to both historical data and retention rates of the control group. The third order outcome will be graduation rates of female STEM majors, compared to both historical graduation rates and graduation rates of the control group.

Procedure

This program will use an Applied Research Oriented Evaluation, and select a few sections of introductory STEM courses (from instructors who teach more than one section) to test the use of exams labeled as "gender-fair." Content is standard across sections of these courses, and the researcher will ensure that each section selected for the intervention group has a paired section, taught by the same instructor, that will not take these exams. With having control sections of the course and identical content delivery in the lecture for each pair of sections, the hypothesis that labeling an exam as "gender-fair" will not only have short-term benefits (higher grades by female students), but will also lead to an increase in the number of female students being retained in and graduating from STEM majors.

Research Questions, Variables, and Analysis

The first question that this evaluation seeks to answer is: Do female STEM students perform better on course examinations when the tests are presented as gender-fair? Unit examinations scores and overall course scores will be used. These measures will be compared across sections and will be written by STEM faculty. An ANCOVA will be used to analyze the results.

The second question this evaluation seeks to answer is: Do female students who continue to take these gender-fair tests persist longer in their STEM majors? This question will be answered by measuring the number of female students who were still in their STEM major the year after the gender-fair tests were used. These data will be compared to both the peer control group as well as historical rates of persistence. An ANCOVA will be used to analyze the data.

A third question this evaluation seeks to answer is: Do female students who take gender-fair examinations retain or strengthen their STEM domain identities? STEM domain identity will be measured both at the beginning of the semester as well as the end of the semester where the gender-fair examinations were used. Data will be analyzed using a repeated-measures ANCOVA.

The final question this program seeks to answer is: Will use of STEM course examinations labeled as gender-fair lead to higher graduation rates of female STEM majors? Graduation rates will be measured using historical data on graduation rates, as well as graduation rates of both intervention and control groups. These data will be analyzed using an ANCOVA.

Program Components

The objective of the Gender-Fair Exam Program is to determine if using course examinations labeled as "gender-fair" will lead to better performance in introductory STEM courses and improved retention and graduation rates in STEM majors (see Diagram 1). The intended first order outcome is that female student performance on examinations will improve after using these gender-fair tests, relative to a control group of students that will take the normal tests. The second order outcomes are that female students will persist longer in their STEM majors (determined if they are still enrolled in their major during the year following the program). The third outcome is that there will be an improvement in overall graduation rates of female STEM majors after the introduction of these gender-fair tests. The following inputs are needed for the program to operate: (1) STEM faculty; (2) Introductory STEM courses; (3) STEM students; (4)

classrooms; (5) STEM exams to be re-labeled as “gender-fair. The program is constrained by the amount of time and money it would take to reissue the examinations with the new label and statement.

The gender-fair exam program for STEM majors will provide female students (see Diagram 2):

(1) Instruction on STEM course content – Direct and indirect instruction will come from faculty in the STEM fields

(2) Constant or improved STEM domain identity – By labeling the examinations as gender-fair, this will buffer from stereotype threat effects and the subsequent chance of disidentification with the STEM domain.

Students will take course examinations, which provide opportunities to solve STEM problems and evaluate their understanding of course concepts (see Diagram 3). Use of gender-fair examinations will allow for a reduction in stereotype threat, which will allow female students to perform better on the course examinations. This will lead to an increased confidence in mastery of the course content and thus will either improve or keep constant their STEM domain identity.

DIAGRAM 1

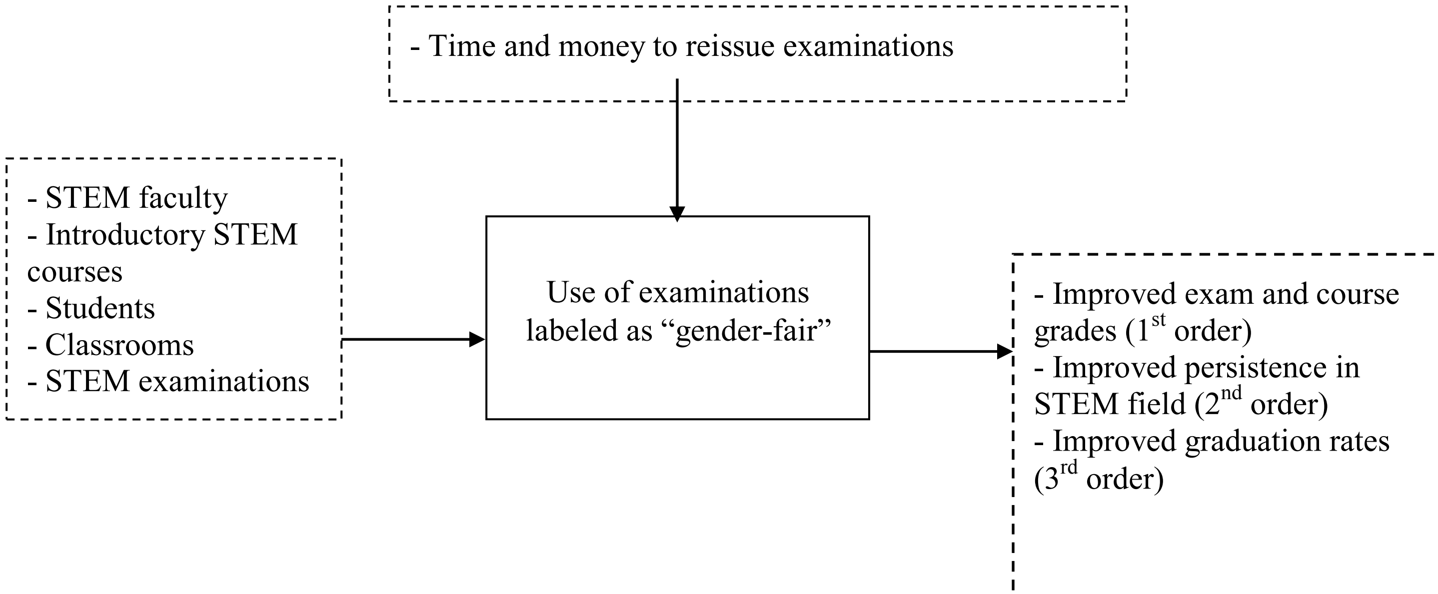

- Time and money to reissue examinations
- STEM faculty
- Introductory STEM courses
- Students
- Classrooms
- STEM examinations
Use of examinations labeled as "gender-fair"
- Improved exam and course grades (1st order)
- Improved persistence in STEM field (2nd order)
- Improved graduation rates (3rd order)

DIAGRAM 2

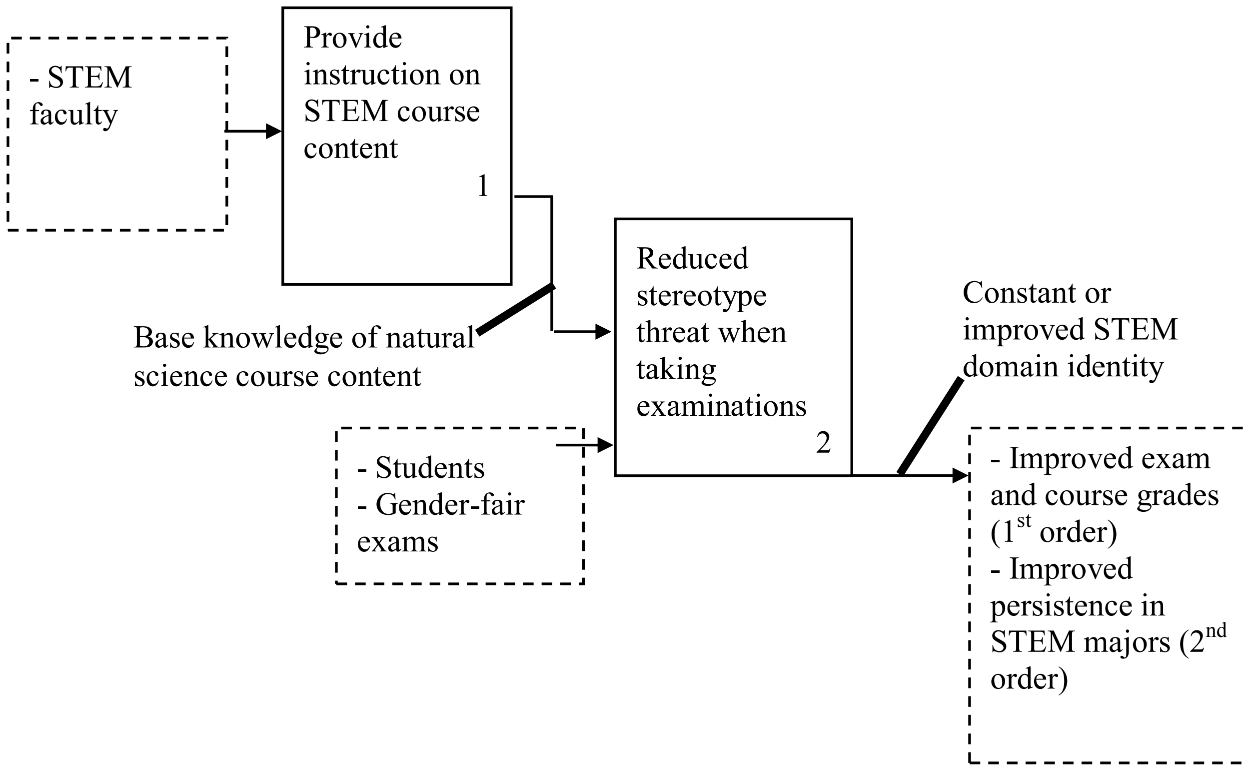
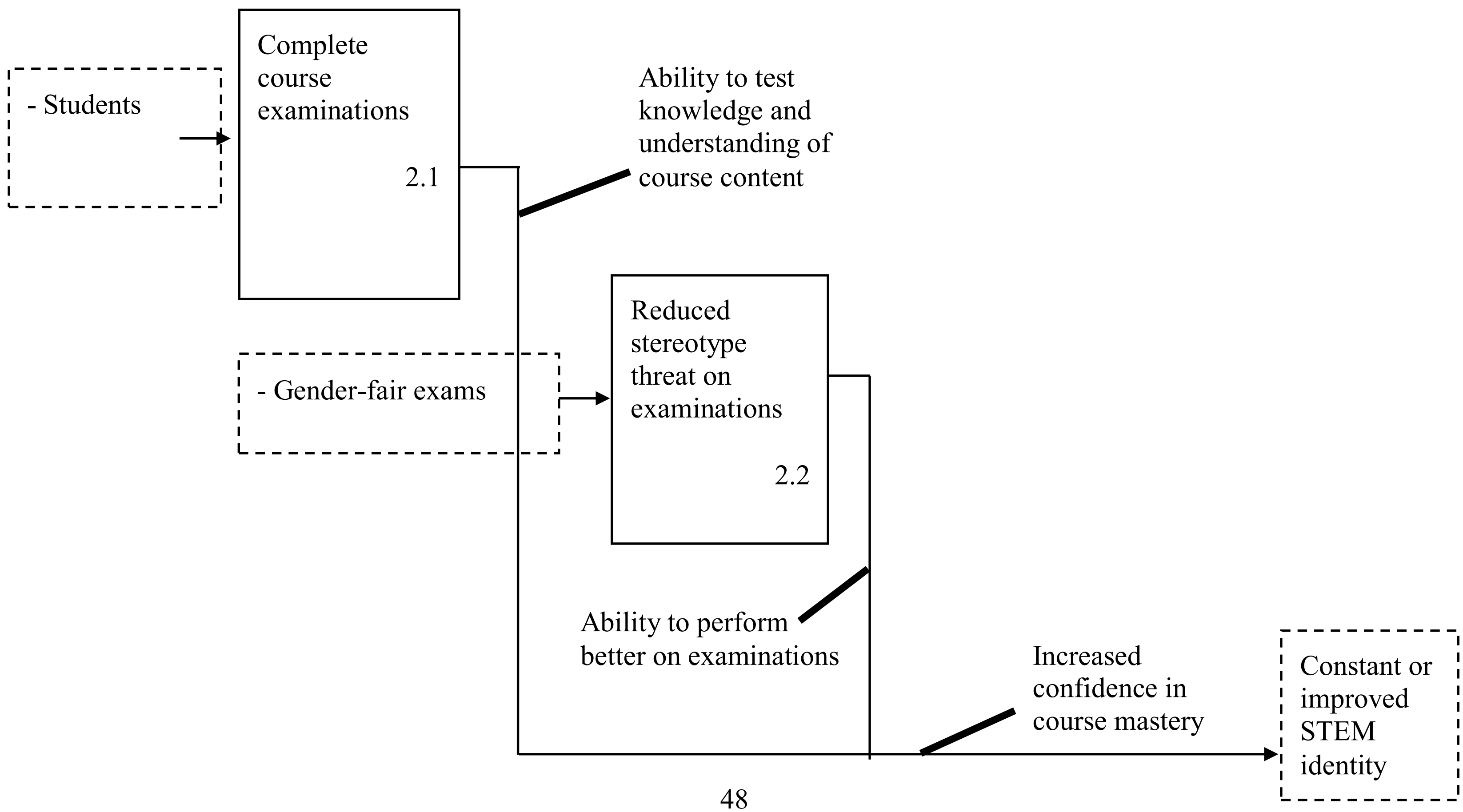

- STEM faculty
Provide instruction on STEM course content
1
Base knowledge of natural science course content
Reduced stereotype threat when taking examinations
2
- Students
- Gender-fair exams
Constant or improved STEM domain identity
- Improved exam and course grades (1st order)
- Improved persistence in STEM majors (2nd order)

DIAGRAM 3– Get practice, immediate feedback, and adaptive and scaffolded assistance



References

American Psychological Association, Task Force on Resilience and Strength in Black Children and Adolescents (2008). *Resilience in African American children and adolescents: A vision for optimal development.* Washington, DC: Author. Retrieved from http://www.apa.org/pi/cyf/resilience.html

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

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype priming on action. *Journal of Personality and Social Psychology,71*, 230-244.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology, 43*, 437-449.

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General*, *136*, 256-276.

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality & Social Psychology*, *71*, 83-93.

Cheryan, S., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of 'model minority' status. *Psychological Science*, *11*, 399-402.

Cokley, K. O. (2002). Ethnicity, gender and academic self-concept: A preliminary examination of academic disidentification and implications for psychologists. *Cultural Diversity and Ethnic Minority Psychology*, *8*(4), 378-388.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, *74*, 865-877.

Dweck, C. S. (1999). Self-Theories: Their role in motivation, personality and development. Philadelphia: Taylor and Francis/Psychology Press.

Ethier, K. & Deaux, K. (1990). Hispanics in Ivy: Assessing identity and perceived threat. *Sex Roles*, 22, 427-440.

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

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.

Hogg, M. A. (2006). Social identity theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (pp. 111-136). Stanford, CA: Stanford University Press.

Inzlicht & Schmader (2011). Introduction. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 544-564.

Johns, M. J., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, *137*, 691-705.

Johns, M., & Schmader, T. (2010). Meta-cognitive regulation as a reaction to the uncertainty of stereotype threat. In R. M. Arkin, K. C. Oleson, & P. J. Carroll (Eds.), The uncertain self: A handbook of perspectives from social and personality psychology (pp. 176-192). Lawrence Erlbaum Associates.

Katz, I., Roberts, S. O, & Robinson, J. M. (1965). Effects of task difficulty, race of administrator, and instructions on digit-symbol performance of Negroes. *Journal of Personality and Social Psychology, 2,* 53-59.

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.

Kray, L. J., Thompson, L., & Gallinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, *80*(6), 942-958.

Levy, B. (1996) Improving memory in old age through implicit self stereotyping. *Journal of Personality and Social Psychology*, *71*, 1092-1107.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*, 302-318.

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24*, 34-50.

Markus, H. R., & Kitayama, S. (1991). Culture and the Self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253.

Marx, (2011). Differentiating theories: A comparison of stereotype threat and stereotype priming effects. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Marx, D. M., & Stapel, D. (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. *Journal of Personality and Social Psychology*, *91*, 243-254.

Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, *44*, 1187-1193.

McFarland, L. A., Kemp, C. F., Viera, L., Jr., & Odin, E. P. (2003, April). *Stereotype threat and male–female differences in test performance.* Paper presented at the 18th annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Murphy, M. C., & Jones Taylor, V., (2011). The role of situational cues in signaling and maintaining stereotype threat. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314-1334.

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). Stereotyping and social reality. Oxford, United Kingdom: Blackwell.

Picho, K., Rodriguez, A., & Finnie, L. (2013). Exploring the Moderating Role of Context on the Mathematics Performance of Females Under Stereotype Threat: A Meta-Analysis, *Journal of Social Psychology*, *153*(3), 299-333.

Rydell, R., McConnell, A., Beilock, S. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology, 96*, 949-966.

Schmader, T., & Beilock, S. (2011). An integration of processes that underlie stereotype threat. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336-356.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83.

Shih, M., (2004). Positive Stigma: Examining Resilience and Empowerment in Overcoming Stigma. *Annals of the American Academy of Political and Social Science*, *591*, 175-185.

Shih, M., Pittinksky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80-83.

Shih, M., Pittinsky, T., & Ho, G. C. (2011). Stereotype boost: Positive outcomes from the activation of positive stereotypes. . In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Spears, R, Gordijn, E., Dijksterhuis, A., & Stapel, D. A. (2004). Reaction in action: Intergroup contrast in automatic behavior. *Personality and Social Psychology Bulletin*, *30*, 605-616.

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.

Srull, T. K., & Wyer, R. S. Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. Journal of Personality and Social Psychology, 37, 1660–1667.

Steele, J. R., & Ambady, N. (2006). "Math is hard!" The effect of gender priming on women's attitudes. Journal of Experimental Social Psychology, 42, 428-436.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. Journal of Personality and Social Psychology, 69(5), 797-811.

Steele, Claude M. (1997). A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. *American Psychologist* 52 (6): 613–629.

Steele, J., James, J. B., & Barnett, R. C. (2002). Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychology of Women Quarterly, 26*(1), 46-50.

Steele, C.M. (2011). Extending and applying stereotype threat research: A brief essay. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*. New York: Oxford University Press.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology, 77,* 1213-1227.

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44,* 445-452.

Tajfel, H., & Turner, J. C. (1979). An intergrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.),*Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454–463.

United States Department of Education. (2012). *Gender Equity in Education: A Data Snapshot*. Retrieved from: http://www2.ed.gov/about/offices/list/ocr/docs/gender-equity-in-education.pdf

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456-467. Wheeler, S. C., DeMarre, K. G., & Petty, R. E. (2004). The roles of the self in priming-to-behavior effects. In A. Tesser, J. Wood, & D. Stapel (Eds.), *The psychology of self*. New York: Psychology Press.

Wylie, P. (1979). *The self-concept: Theory and research on selected topics*. Lincoln, NE: University of Nebraska Press.