

The Pennsylvania State University

The Graduate School

College of Education

CAN A SHORT VIDEO MODULE TEACH STEREOTYPE THREAT AND METHODS OF
PREVENTION TO WHITE, WOMEN, PRE-SERVICE TEACHERS?

A Dissertation in

School Psychology

by

Chloe Y. H. Barry

© 2017 Chloe Y. H. Barry

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2017

The dissertation of Chloe Y. H. Barry was reviewed and approved* by the following:

Beverly J. Vandiver
Associate Professor Emerita of Education
Dissertation Advisor
Chair of Committee

Barbara A. Schaefer
Associate Professor of Education

Deborah F. Atwater
Associate Professor Emerita of Communication Arts and Sciences, and African and African-American Studies

Dawn P. Witherspoon
Associate Professor of Psychology

James C. DiPerna
Professor of Education
Professor in Charge of School Psychology

*Signatures are on file in the Graduate School.

ABSTRACT

The purpose of the study was to examine whether a brief video module on stereotype threat (ST) could be used to educate pre-service teachers about its effect and ways to prevent such effect on elementary children. This study was intended to be a first step in helping teachers learn about ST, so they can reduce the impact of ST on their students. Participants were 60 White women, who were pre-service elementary teachers from a predominantly, White university in the Mid-Atlantic region of the United States. The research design was a 2 x 2 between subjects experimental design, which involved two manipulations. The first manipulation was the module (exposure or not) and the second manipulation was hypothetical scenario (ST or No-ST). Half of the participants were randomly exposed to a brief ST video module, which described ST, its effects, and ways to prevent such effects, and the other half did not receive the module. All participants were asked to complete a knowledge measure about classroom organization. Then the same participants were randomly assigned to a scenario of a student experiencing ST or math anxiety. Participants rated the likelihood of using a list of various interventions, which had been determined to be ineffective as well as effective ST strategies. They also completed a knowledge measure on ST. A series of ANOVAs and MANOVAs, followed with descriptive discriminant analyses were run to test the two hypotheses. Neither of the hypotheses were supported: Exposure to the ST video module did not increase the likelihood of the pre-service teachers to indicate use of strategies for the ST scenario more than those in the control condition (no module) exposed to the same scenario. However, a main effect was found for scenario. Regardless of exposure to the ST video module or not, participants exposed to the ST scenario were more likely to use the strategy of explaining the anxiety than those participants exposed to the No-ST scenario. This could mean that there may be a confounding variable and that

participants not exposed to the ST video module interpreted the ST scenario as reflective of another theoretical base, like test anxiety. It is possible that more pilot work was needed on the ST video module—balancing the information on identifying ST and selecting the most effective strategies. Furthermore, the item difficulty of the ST knowledge needed to be increased to discriminate between those exposed to the ST video module versus not. And the approach to rating the intervention strategies increased the rating of most interventions as likely to be used. Beyond these changes, replicating this study with larger and more diverse samples is recommended.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	ix
ACKNOWLEDGMENTS	x
DEDICATION	xii
INTRODUCTION	1
LITERATURE REVIEW	5
Overview of ST	5
Seminal ST Research	6
The Basic Elements of ST Research Design	11
The Processes of ST	14
Physiological stress on working memory	15
Self-monitoring	15
Thought suppression	16
Psychological Interventions of ST	16
Prevention of ST	17
Explaining the task differently	17
Making self-affirmation	21
Providing a role model	26
Switching identities	30
Viewing the brain as a muscle	33
Providing external attributions	36
Summary	38
Wise Interventions	39
Teacher Change	41
The Current Study	43
METHOD	44
Pilot 1	44
Research Design	44
ST video module	44
Scenarios	45
Participants	47
Measures	48
Demographic information	48
Test of Application	48
Test of Knowledge	49
Procedures	50
Results	51
Other Revisions to Research Design	51

Pilot 2	52
Research Design	52
Participants	52
Measures	52
Procedures	53
Results	53
Other Revisions to Research Design	55
Actual Study	56
Research Design	56
Knowledge Module	56
Scenarios	57
Data Management of the Sample	58
Participants	58
Measures	58
Demographic information	60
Classroom organization	60
Test of Application	60
Test of Knowledge	61
Procedures	61
RESULTS	63
Preliminary Analyses	63
EFA	63
Demographic variables and research design	66
Descriptive Statistics	69
Hypothesis 1	70
Hypotheses 2	72
DISCUSSION	77
Hypothesis 1: Exposure to ST Video Module and Knowledge	77
Hypotheses 2: ST Video Module Exposure and Selected Prevention Strategies	79
Limitations	81
Future Direction	83
Implications	85
Conclusion	88
REFERENCES	89
APPENDIX A: ST Video Module Transcript	104
APPENDIX B: ST TOA in Pilot 1, 2, & Actual Study	106
APPENDIX C: No-ST TOA in Pilot 1, 2, & Actual Study	109
APPENDIX D: Classroom Organization in No-ST TOA in Pilot 1, 2, & Actual Study	112

APPENDIX E: TOK in Pilot 1, 2, & Actual Study	113
APPENDIX F: Recruitment Email	116
APPENDIX G: Correlations for TOK and TOA	117
APPENDIX H: Summary of EFA, Maximum Likelihood Extraction, Promax Rotation of TOA	118
APPENDIX I: Frequency Table of Response Rates for TOK and Classroom Organization ($N = 60$)	119

LIST OF TABLES

TABLE 1: Frequencies of Participants' Professional Development on a Given Topic	59
TABLE 2: Percentage of Academic Class Standing based on the Conditions of Module and Scenario	67
TABLE 3: Percentage of Family Income based on Conditions of the Module and Scenario	68
TABLE 4: Means, Standard Deviations, Skew and Kurtosis of the TOK and TOA.....	69
TABLE 5: Summary of ANOVAs based on Module and Scenario on Classroom Organization and Three TOK Total Scores ($N = 60$)	71
TABLE 6: Summary of MANOVA of Module Exposure and Scenario Conditions on Explain Anxiety, Math Puzzles, and Remind Student ($N = 60$)	73
TABLE 7: Structure Coefficients and Standardized Function Coefficients for Descriptive Discriminant Analyses for the ST Prevention Strategies Based on the ST Scenario ($N = 60$)	74
TABLE 8: Summary of MANOVA of Module and Scenario Conditions on Generalized Prevention	75
TABLE 9: Summary of MANOVA based on Module and Scenario for the Set of Ineffective Strategies ($N=60$).....	75
TABLE 10: Summary of One-Way ANOVA based on Module and Scenario Conditions on the TOA Item Take No Action ($N = 60$).....	76

LIST OF FIGURES

FIGURE 1: Graphical depiction of the centroids for the discriminant function of the ST prevention strategies ($N = 60$).....	74
--	----

ACKNOWLEDGEMENTS

This dissertation and really my entire graduate school career would not have been possible without my advisor and dissertation chair, Dr. Beverly Vandiver. I am so thankful that you saw something in me all those years ago. Thank you for all of your hard work, expertise, copious amounts of feedback, and encouragement. You have helped me grow into a much stronger writer, psychologist, and person. I would also like to thank my dissertation committee members, Dr. Deborah Atwater, Dr. Barbara Schaefer, and Dr. Dawn Witherspoon. I know my dissertation journey has been a long one and I thank you for your patience, flexibility, and understanding throughout this process. I am so grateful for all your contributions to making this dissertation what it is through your guidance and insightful feedback. Baby cohort (Arlene Ortiz, Leah Kelm, Brianne Mintern, Sarah Wollersheim Shervey, and Emily Perlman), thank you for being my group of friends over these past years. You made me feel welcome and gave me a rare sense of belonging.

A big thank you to my family. My parents, Jill and Joe Barry, worked tirelessly to ensure that my sisters and I had the tools to succeed. Thank you for creating an environment where education was a top priority. Leigh and Ferin Barry, thank you for always being there when I needed and for inspiring me to strive to be a better example for both of you and for others. Also thank you to my extended family who throughout this process have made many calls, sent many texts, and many emails with encouraging words and inquiries as to when I would be graduating. I have appreciated your unfaltering persistence on coming to my graduation. I would also like to thank Dr. Michael Stoyer. Thank you for your support, your never ending offers to help me in this process, and your flexibility in picking up the slack whenever my world turned to dissertation. I know it has not always been easy for you, and I have no adequate words to

describe how grateful I am that I had you to help me (and sometimes push me) through this whole process.

Lastly, I would like to acknowledge some of the people integral to laying the brickwork for my path towards graduate school. Students are blessed to have one exceptional teacher during their formative years, but I was lucky to have several. Thank you Gail Aita, Patrick Gallagher, Jason Smith, and Tina Tobiason for fostering and supporting my passion for learning.

DEDICATION

This dissertation is dedicated to my family. Mahalo.

INTRODUCTION

By the age of five, almost everyone within a culture is aware of certain racial and ethnic stereotypes (Aronson, 2009). Young children are susceptible to internalizing these learned stereotypes and prejudices (Martines, 2008), meaning that from a young age stereotypes can be constructed, maintained, and internalized. For children, the development of racial/ethnic stereotypes has many implications. Stereotypes can affect students' social groups, which may exclude them from school enrichment and job opportunities (Steele, 2010). In addition, when students are required to cope with stigmas related to stereotypes, it can deplete their intellectual resources, which, in turn, may impair their decision-making skills (Inzlicht, Tullett, Legault, & Kang, 2011) and attention (Blair & Steele, 2010).

In 1968, Jane Elliot split her third grade class into two groups based on eye color (blue-eyed and brown-eyed students) to teach her students about discrimination (Peters, 1971). On alternate days, children with one eye color were informed that they were superior to students with the other eye color. For example, the blue-eyed students were given the privilege to sit in the front of the classroom, to stand first in line for lunch, and to exclude brown-eyed students from playing with them. Within the day, the blue-eyed students were mean to their brown-eyed peers and performed better than usual in math and reading. In contrast, the brown-eyed students became timid and their performance on assignments suffered. On the next day, this process was reversed. The same behavior observed in the blue-eyed children occurred in the brown-eyed children. This experiment conducted in a classroom setting showed that teachers, whether purposeful or not, can influence students' development of stereotypes and the effect of promoting such stereotypes.

It is imperative that teachers are aware of invoking or reinforcing stereotypes within the classroom setting, because students can internalize stereotypes as shown in the Blue Eye/Brown Eye Experiment and be affected by them, as evidenced by the lowered scores achieved by the students when subjugated to stereotypes. The Blue Eye/Brown Eye Experiment serves as an exemplar of stereotype threat (ST), which explains the effect of stereotypes on individuals' performance, including students' academic performance (Steele & Aronson, 1995).

ST is the awareness that one's actions may confirm a stereotype about one's group (Steele & Aronson, 1995). Steele (1997) believed ST may also affect performance. Initially, the fear and thoughts related to confirming a stereotype interfere with performance. Then ST affects performance by becoming chronic. Chronic ST occurs in competitive situations in which the person is consistently in an out-group environment.

Chronic ST could cause disidentification from the domain related to the stereotype (Griffin, 2002; Steele, 1997). For instance, if a student frequently experiences ST during math tasks, the student may no longer be interested in mathematics even though the student has the potential to do well. Disidentification is considered to be a reconceptualization of the self that no longer includes the domain related to the threatening stereotype. Although this new lack of interest in math can be protective, it might have an effect on the representation of groups in certain professions. For instance, domain abandonment due to chronic ST may be one of the reasons why women are under-represented in math-related fields. A woman interested in pursuing a math-related field may eventually reconceptualize herself to no longer identify math as important to protect herself against the feelings of ST and pressures related to being in a male-oriented math environment. Disidentification from academics at an individual level can be costly (i.e., Steele, 1997; Steele, Spencer, & Aronson, 2002). For instance, a loss of interest in

academics due to disidentification can cause students to drop out which is related to less average lifetime income (Watts, 2001) and the increased likelihood of poor health later in life (Ross & Wu, 1996). Therefore, it is important to find effective ways to prevent students from experiencing ST. Most research on the prevention of ST has been conducted in academic settings at the college level, but by this time, the effects of chronic ST leading to disidentification of a domain may have already taken place. For example, the effects of chronic ST may have occurred among high school girls who were found to opt out of classes that involved the application of higher-level mathematics skills more often than the boys (Riegle-Crumb, 2007). Knowing that chronic ST can result in disidentification in mathematics or academics, in general, makes ST research on interventions for teachers essential.

Since Steele and Aronson (1995) tested the concept of ST, hundreds of studies have been conducted on ST in various disciplines. Rios, Cheng, Totton, and Shariff (2015) examined how ST affected Christians on science achievement. In another recent ST study, Najdowski, Bottoms, and Goff (2015) studied ST in Black men and its effect on their encounters with police officers. Although ST can be experienced by anyone who identifies with a group who is negatively stereotyped (Steele, 1997), the body of research on ST has focused on gender and ethnic-based stereotypes. Thus, the general consensus is that ST has a negative effect on ethnic minorities and women. In support of this idea, Nguyen and Ryan (2008) conducted a meta-analysis consisting of 116 research studies and examined the effects of ST on both gender- and race-based stereotypes. Overall, Nguyen and Ryan found that ST negatively affected both ethnic minorities and women.

Understanding that women and ethnic minorities experience ST is important; however, a need also exists to develop methods to reduce ST in individuals, especially school age youth. As

stated earlier, children begin internalizing stereotypes at a young age (Martines, 2008). Teachers need to be aware of invoking or reinforcing stereotypes in the classroom setting, because students may internalize stereotypes and be affected by them, as evidenced in the Blue Eye/Brown Eye Experiment. It is important for teachers to have an awareness of ST, because ST may be interfering with the performance of their students. Therefore, it is important to find effective ways for teachers to learn about ST and to develop techniques to prevent students from experiencing ST.

At this time, the focus of research has not been on educating teachers about the effects of ST and the use of effective methods to reduce students' development of chronic ST. Thus, the purpose of this study was to examine whether educating pre-service teachers about ST and possible effective interventions via a short video module would improve their ability to identify and respond appropriately to ST. This experiment is considered a first step toward an applied approach to the prevention of ST and may lead to increasing teachers' involvement in preventing ST in the classroom setting.

LITERATURE REVIEW

The purpose of this study was to examine whether providing pre-service teachers with information about ST would increase their skill in perceiving and responding to children, who may experience such a threat. To date, this area, training teachers to reduce ST in children, has not been a focus of research. As a result, the literature review is focused on ST and prevention research. First, an overview of ST is provided, including defining the concept and the situations in which ST occurs. In subsequent sections, a summary and critique are provided about the Steele and Aronson's (1995) seminal research on ST. Furthermore, an overview of the basic ST research design is provided as well as the processes considered to result in the decline of performance. The literature review ends with a summary of strategies that have been empirically supported in preventing or reducing the occurrence of ST. Then, there is a discussion on the timing of interventions and teacher change. Finally, the rationale for the current study, guiding research questions, and hypotheses are provided.

Overview of ST

Steele and Aronson (1995) developed the theory of ST to explain how a stereotype may have an impact on someone's performance in a challenging situation. Individuals experience ST when they (a) are invested in an area of performance, (b) are conscious of a stereotype (negative or positive), and (c) hold a belief that the stereotype could be confirmed as true about their group and themselves through their actions. The perceived threat could occur due to individuals believing the stereotype or being concerned about others having the stereotyped perception about the threatened group (Shapiro & Neuberg, 2007). ST can be enacted through different kinds of stereotypes and in different situations. This threat has been activated through various stereotypes, such as race (Steele & Aronson, 1995), gender (Nguyen & Ryan, 2008), and age

(Hess, Auman, Colcombe, & Rahhal, 2003), resulting in decreased performances in various domains like memory (Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005), academics (Steele & Aronson, 1995), driving (Yeung & von Hippel, 2008), and work place (Bergeron, Block, & Echtenkamp, 2006). The origin of ST is considered to have come from many different sources, all of which focused on an awareness of stereotypes and its linkage to anxiety (Allport, 1954; Carter, 1991; Goffman, 1963; Steele, 1990; as cited in Steele & Aronson, 1995). Steele's linkage of stereotype to anxiety led to the latest iteration. Steele (1990) contended that because of society's negative images of an African American's ability, these racial/ethnic minority students are likely to internalize an anxiety of inferiority, and that the anxiety could be triggered by cues in the environment. Based on this premise, Steele and Aronson (1995) conducted the first research on ST, hypothesizing that the Black-White achievement gap was not entirely due to differences in education, but that ST played a role.

Seminal ST Research

In their seminal ST research, Steele and Aronson (1995) conducted a series of four studies. In Study 1, Black and White, college students ($N = 114$; sample size by race was not provided) were administered a 30-minute test, which contained difficult items from the verbal portion of the Graduate Record Examination (GRE; <http://www.ets.org/gre>). Then participants answered questions about academic confidence, self-worth, anxiety, bias level of the test, perceived performance, and the experimental manipulation. The design was a 2 (race) x 3 (condition) between-group design, in which Black and White participants were randomly assigned to one of three conditions (ST, No-ST, and Challenge). In the ST condition, the 30-minute test was described to the students as a diagnostic test of intellectual ability. Steele and Aronson (1995) believed calling the test diagnostic would activate a racial stereotype about

intellectual ability for the Black participants. For the No-ST condition, the test was called a problem-solving task. And in the Challenge condition, the task was introduced to the students as a challenge. It was hypothesized that the Challenge condition would increase the participants' motivation to perform well on the test.

A 2 (race) x 3 (condition) ANCOVA was conducted, with self-reported SAT scores as the covariate and current verbal GRE score as the dependent variable. Adjusting for SAT scores, main effects were found for both race and condition. In general, White students scored higher than the Black students on the verbal GRE. Also, students in the Challenge condition scored higher on the verbal GRE than students in the ST and No-ST conditions. No race x condition interaction was found for test performance. Steele and Aronson (1995) noted that the lack of an interaction effect was probably due to the scoring pattern based on race in the Challenge condition. White students scored slightly higher than Black students in the Challenge condition, which was believed to have undermined the interaction effect. Thus, Steele and Aronson conducted Bonferroni contrasts of conditions by each race while controlling for SAT scores. As predicted, Black students in the ST condition underperformed in comparison to Blacks in the No-ST and Challenge conditions. Also, Black students in the ST condition performed worse than White students in the ST condition. Steele and Aronson concluded that when a test is described as diagnostic of ability, ST appears to be invoked and results in depressing the scores of Black students. Finally, no statistically significant differences were found for race or condition on self-rated academic competence or self-worth.

In the first experiment, Steele and Aronson (1995) also performed a 2 (race) x 3 (condition) ANCOVA, with previous SAT scores as the covariate and self-reported performance comparison to others as the dependent variable. A main effect was found for condition; Black

participants in the ST condition perceived their performance as lower than those in the No-ST and Challenge conditions. Steele and Aronson (1995) did not provide a rationale for why perceived performance in comparison to others was lower for the Black students in the ST condition in comparison to Black students in the No-ST condition.

In the second experiment, Steele and Aronson (1995) tested whether the fear of confirming a negative stereotype could be detected by a general anxiety measure. Participants were 20 Black and 20 White, female, undergraduate students. The general set-up for this study was similar to Study 1 in that the ST and No-ST condition were used, but the Challenge condition was not used, resulting in a 2 (race) x 2 (condition) design. In contrast to the Study 1, the task was shortened by three items, students were only allowed 25 minutes to complete the task instead of 30, the task was administered on a computer, and the Spielberger State Anxiety Inventory (Spielberger, Gorsuch, Lushene, & Jacobs, 1983) was given to the participants after the test was completed.

The 2 (race) x 2 (condition) ANCOVA, with self-reported SAT scores as the covariate and self-reported anxiety as the dependent variable, was not statistically significant, but a similar ANCOVA with test performance as the dependent variable was statistically significant for the interaction term as well as for the main effect for race. Based on planned contrasts, Black and White students were found to perform about equally well in the No-ST condition, but Black students performed worse than White students in the ST condition. The race x condition interaction found in Study 2 but not in Study 1 may have been due to differences in test composition and the length of the test. Steele and Aronson (1995) noted that ST slows and decreases the accuracy of students' work. Thus, Study 2 showed the effect of ST more

dramatically because the participants in Study 2 encountered more multistep items at the beginning of the test and had 5 minutes less than those in Study 1 to complete the test.

To extend what was learned from the first 2 studies, in Study 3, Steele and Aronson (1995) tested whether the fear of confirming a stereotype would lead to students disassociating from the stereotyped group. Word activation tasks (stereotype and self-doubt) were added. It was expected that students experiencing ST would have more stereotype and self-doubt related thoughts, which would result in elevated scores on the activation tasks. The sample was 35 Black and 33 White undergraduates, who were randomly assigned to one of three experimental conditions (ST, No-ST, and Control). The task in the ST condition was described as diagnostic of ability, and the task in the No-ST condition was described as non-evaluative. In the Control condition, no explanation of the task was given. After the task introduction, all participants completed questions on demographics, and completed a stereotype and self-doubt activation word task, which was introduced as a visual and recognition task of words. On the word task, students were asked to complete words that contained at least two letters blanked out. The self-doubt activation had students to complete words such as loser and weak. The stereotype activation task consisted of words such as minority and Black. Finally, the participants were given items from the GRE to complete.

Steele and Aronson (1995) reported a statistically significant ANCOVA, with SAT scores as the covariate, for the interaction between race and condition on the stereotype activation task. Additionally, the main effects of race and condition were also found to be statistically significant on stereotype activation. Contrast analyses indicated that Black students in the ST condition correctly completed more words related to race than Black students in the No-ST and Control conditions, and White students in the ST, Control, and No-ST conditions.

An ANCOVA was also conducted, with self-reported SAT scores as the covariate and self-doubt as the dependent variable. The interaction of race x condition was statistically significant. Contrast analyses revealed that Black students in the ST condition had more self-doubt completions in comparison to Black students in the No-ST and Control conditions and White students in ST, Control and No-ST conditions.

Steele and Aronson (1995) also examined whether students reported their race when completing the demographic questions. An ANCOVA, with SAT scores as a covariate, revealed a statistically significant race x condition interaction on reporting race. Approximately 25% of the Black students in the ST condition reported their race. In contrast, all participants in the other groups (White students across all conditions, and Black students in the No-ST and Control conditions) reported their race.

In the final study, Steele and Aronson (1995) had a sample of 23 White and 24 Black undergraduates, who were given the same 25-minute test from the Study 2 and were randomly assigned to one of two conditions (Race Prime and No-Race Prime). In the Race Prime condition, participants had to first complete a demographic section, which contained a question on race. This question on race was not included in the demographic section in the No-Race Prime condition. A 2 (race) x 2 (condition) ANCOVA was performed on test performance, using self-reported SAT scores as a covariate. The interaction was statistically significant. Black students in the Race Prime condition performed worse than the Black students in the No-Race Prime condition, and White students in the Race Prime condition. However, in the No-Race Prime condition, Black students' performance was equivalent to the White students' performance in the Race Prime and No-Race Prime conditions.

Steele and Aronson's (1995) work provided the basic research design to study ST, demonstrating its existence, and stimulating significant research on ST. One of the strengths of Steele and Aronson's research was the replication of the effects of ST across four studies with four different samples. However, a limitation of these studies was the use of the same population—college students. Other limitations of these studies were the lack of reporting effect sizes, and the reliance on self-reported SAT scores, instead of actual SAT scores.

The Basic Elements of ST Research Design

To aid in the review of the relevant research on ST, the basic research design used in the majority of these studies is presented. Generally, the research design is similar to the one used by Steele and Aronson (1995), in which several criteria undergird the design. One, to enact ST, a social identity associated with a specific stereotype must be selected. For instance, Steele and Aronson (1995) used a racial stereotype relevant to African Americans. ST can be activated in any group that has a stereotype whether it is positive or negative. Racial/ethnic minorities (Alter, Aronson, Darley, Rodriguez, & Ruble, 2009; Steele & Aronson, 1995) and women (Hermann & Vollmeyer, 2016; Inzlicht & Ben-Zeev, 2000; Schmader, 2002) have been the populations most commonly used in ST research, but other groups, like low socioeconomic status groups (Croizet & Claire, 1998), the elderly (Hess, Hinson, & Hodges, 2009; Levy, 1996), and athletes (Yopyk & Prentice, 2005) have also been examined. Two, a pre-existing relationship is considered to exist between the identity and the domain under study. For example, men have long been considered better at mathematics than women or Whites have been considered better at academics than Blacks. Three, the domain must be important to the participants. Thus, women must be invested in doing well in mathematics and Blacks must be invested in doing well academically. Domain investment or identification plays a significant role in whether individuals experience ST. An

assumption of ST theory is that for success to be sustained in a domain a person must identify with it, which causes them to be self-evaluative and accountable (Steele, 1997). This accountability creates a situation in which students are more susceptible to ST, because they are worried about their domain performance in terms of self-evaluations, and the perceptions and the evaluations of others of them. Additionally, participants need to identify with the threatened social identity in order to experience ST. Schmader (2002) found that women who identified highly with math, but not with their gender were less likely to experience ST than women who were both highly math and gender identified. Thus, ST is thought to have an impact on participants' task performance when participants identify with a threatened social identity (Schmader, 2002) and with the task domain (Steele & Aronson, 1995).

Typically, an experimental between-groups design has been used to test ST. At minimum, there are two conditions (ST and No-ST). Participants are randomly assigned to the conditions. In the ST condition, various procedures have been used to activate ST. Common methods to activate ST have involved the following: (a) describing the task as diagnostic (Steele & Aronson, 1995), (b) invoking minority status through being the only member of a group in the room (Inzlicht & Ben-Zeev, 2000), (c) describing or introducing a stereotyped belief (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004), and (d) asking students to report information about the stereotyped identity to make it more salient (McGlone & Aronson, 2006).

The No-ST condition may involve the absence of the method used to activate ST. McGlone and Aronson (2006) in the No-ST condition had students report their demographic information after the test to prevent the demographic items from making a stereotyped identity salient. For the No-ST condition, Inzlicht and Ben-Zeev (2000) asked women to complete a task in the presence of other women. In contrast, women completed the task in the presence of men

in the ST condition. In other studies, the No-ST condition has been designed to do the opposite of what was done in the ST condition. For instance, Spencer, Steele, and Quinn (1999) told the participants in the No-ST group that no gender differences had been found on the test, but in the ST condition participants were told that gender differences had been found.

In some studies, additional experimental conditions have been used. Steele and Aronson (1995) included a Challenge condition in their seminal research. Levy (1996) manipulated the type of attribution associated with a successful performance: one allowed participants an external attribution of successful performance, and another allowed participants an internal attribution of successful performance. Both Levy (1996), and Steele and Aronson (1995) decided to drop the additional conditions in subsequent studies, after finding little benefit to adding additional conditions.

After the period in which a stereotype has been activated or not, the participants are asked to complete a task. In most ST research, the task has been related to academic achievement or cognitive ability. One of the most common academic tasks has been drawn from either the Graduate Record Examination (GRE; <http://www.ets.org/gre>) or the SAT (<http://www.sat.collegeboard.org/>). Participants' scores on task performance are used to test for the presence of ST. Following the task, additional measures are administered to the participants, with the goal of allowing the researchers to examine the presence of possible moderators (e.g., education level and stigma consciousness; Hess, Hinson, & Hodges, 2009) and other factors (e.g., test anxiety; Steele & Aronson, 1995; worry; Hermann & Vollmeyer, 2015) that may be associated with the ST or No-ST conditions. Steele and Aronson (1995) had students complete an anxiety measure after completing the academic task in their second study to examine whether the fear of confirming a stereotype caused by ST could be detected in a state anxiety measure.

ST research may also contain a manipulation check. Steele and Aronson (1995) at the end of their first study asked all the participants about the purpose of the study. The manipulation check was to ensure that the ST group was most likely to believe that the study was about evaluating abilities in comparison to the Challenge and the No-ST groups. Although Steele and Aronson included a manipulation check, they only reported its use in one of the four studies. Many studies on ST have not reported using a manipulation check (e.g., Keller, 2007; Spencer et al., 1999; Taylor & Walton, 2011).

ANCOVA has typically been the primary statistical analysis used to test for ST (Gonzales, Blanton, & Williams, 2002; Keller, 2007; Marx, Ko, & Friedman, 2009; Taylor & Walton, 2011). The independent variable is usually the ST manipulation conditions, the dependent variable is usually the scores on the task performance measure, and the covariate has typically been participants' scores from previous task performance measures, such as the SAT. These prior scores have been either obtained directly from the participants or gathered from student records. Planned contrasts have also been used (Croizet, Dutrevis, & Desert, 2002; Martens, Johns, Greenberg, & Schimel, 2006; Stricker & Ward, 2004). Finally, another commonly used method of analysis has been to use ANOVAs, not ANCOVAs (Spencer et al., 1999; Steele & Aronson, 1995, Taylor & Walton, 2011).

The Processes of ST

Specific processes are considered to be associated with the activation of ST, and thus, the lowering of performance. According to Schmader, Johns, and Forbes (2008), three processes are believed to occur when ST is activated and result in the decline of performance. All three of the processes are thought to involve executive resources that are needed to perform well on a task.

The three processes are (a) the effect of physiological stress on working memory, (b) self-monitoring, and (c) thought suppression.

Physiological stress on working memory. Physiological stress response is considered to have a direct role on the decrease in performance when individuals experience ST (Schmader et al., 2008). Physiological stress is thought to reduce prefrontal processing, which can hamper working memory, and different tasks can be more affected than others (Schmader et al., 2008). Beilock, Rydell, and McConnell (2007) found that women who experienced ST were less accurate on difficult math problems that required more information to be held in memory in comparison to women who were in the No-ST condition. Furthermore, Beilock and colleagues (2007) found women had lower performance on math problems (while doing a spatial or verbal task) in the ST condition than those in the control condition. Additionally, no difference was found in completion rates between the verbal and spatial tasks in the No-ST condition. However in the ST condition, verbal tasks were completed at a slower rate in comparison to the spatial tasks. Beilock et al. (2007) concluded that verbal tasks seem to be competing for the same mental resources as ST.

Self-monitoring. The effects of ST may lead students to excessively monitor their performance (Schmader et al., 2008). Students experiencing ST have reported spending more time worrying about and monitoring performance than students not experiencing ST. Beilock and colleagues (2007) believed excessive self-monitoring added more to the students' cognitive load. Forbes, Schmader, and Allen (2008) examined the role self-monitoring may have when external feedback (real or perceived) is given in an academic setting. An event-related potential methodology was used to examine how ST affects ethnic minority students' sensitivity to task errors. Eleven Black and forty-six Latino/a college students completed a task requiring them to

quickly decide whether all the letters or arrows in an array were the same or not. The task was described as either a pattern recognition task (No-ST) or a diagnostic intelligence task (ST). After answering each item, students were given feedback indicating if the answer provided was correct. As expected, racial/ethnic minority students who identified with academics showed a higher brain response to errors in the ST condition than in the No-ST condition. The higher error response in the ST condition was associated with brain activity that implied more error monitoring processes than in the No-ST condition.

Thought suppression. An increase in thought suppression is considered to occur when ST is activated. This mechanism is considered to serve to decrease negative thoughts (Schmader et al., 2008). Research on thought suppression in relation to ST is still in its infancy (Schmader et al., 2008), but previous research has shown that ST can lead to self-doubt (Steele & Aronson, 1995) and task-related worries (Beilock et al., 2007). Schmader and colleagues (2008) posit that the reason previous studies have not shown a consistent relationship between anxiety and ST is because students are actively trying to deny the experience of threat. The process of thought suppression is effortful and is another way that ST may deplete working memory, resulting in impaired task performance (Schmader et al., 2008). Thought suppression as a resource depleting process has been supported in research (Muraven & Baumeister, 2000). In addition, emotional suppression has been found to also increase students' ability to access thoughts related to anxiety (Wegner, Erber, & Zanakos, 1993).

Psychological Interventions of ST

Two approaches to preventing or reducing the effects ST are reviewed. One set of approaches emanates from the specific research on ST. The other set of approaches comes from the emergence of a new class interventions, called wise interventions (Walton, 2014).

Prevention of ST

An area to emerge out of the ST research that started to flourish in 2003 is the prevention of ST. Numerous prevention methods have been studied, both in real world and laboratory settings. Prevention strategies that have received support over multiple studies will be summarized. For clarity and ease of understanding, the review of these studies has been grouped into six themes. The prevention methods reviewed are (a) explaining the task differently, (b) making self-affirmation, (c) providing a role model, (d) switching focus from a threatened identity, (e) viewing the brain as a muscle, and (f) providing external attributions for difficulty.

Explaining the task differently. One strategy used to reduce ST is changing the way the task is described. Steele and Aronson (1995) found that the explanation of a task influenced the activation of ST. Thus, informing the participants that a test was not diagnostic reduced the effects of ST in African American students compared to the African American students who were told that the test was diagnostic of ability.

This method of prevention originates from the Yerkes and Dodson Law (Yerkes & Dodson, 1908). According to Yerkes and Dodson, motivation is inversely related to the difficulty of a task; thus, an easy task could be learned with higher motivation or pressure, but a more difficult task may only be learned with weaker motivation or pressure. Sarason and Palola (1960) added to Yerkes and Dobson's research by examining in college students (a) the relationship between motivation and task difficulty in humans, (b) the role of task description on task completion, and (c) the relationship between task characteristics and anxiety in stressful and threatening situations. An interaction was found between level of anxiety, task description, and task difficulty. On the easy task, those with high anxiety performed better than those with low anxiety, but on the more difficult task, those with low anxiety performed better than those with

high anxiety. Sarason and Palola also found that high anxiety students performed better under a neutral test description than low anxiety students, and high anxiety students performed better than low anxiety students in an intelligence task description condition. Changing a task description seemed to enhance or lessen the anxiety felt by participants during the task. Across three experiments, Sarason and Palola (1960) found an interaction effect between level of anxiety, task description, and task difficulty. They concluded that anxiety level, task difficulty, and task description can affect the performance of participants.

Katz, Roberts, and Robinson (1965) used the Yerkes and Dodson Law (Yerkes & Dodson, 1908) to examine task description and task difficulty. This study is considered to be a test of ST, because of the task difficulty and task description manipulations, both similar to those used by Steele and Aronson (1995). Although the concept of ST had not yet been labeled as such, Katz and colleagues (1965) used White experimenters to elicit more drive, which is similar to other ST research that has used White confederates or experimenters to elicit ST (Marx & Goff, 2005; Stone & McWhinnie, 2008). Participants, Black males ($N = 184$), were randomly assigned to one of three task difficulty (difficult, moderate, or easy) conditions and to one of two task presenters of a different race (Black or White). All participants were told that the task was a motor coordination task. Participants assigned to the most difficult task and had the task presented by a White experimenter completed the task more quickly than when a Black experimenter presented the task. However, the race of the experimenter on the easy and moderate tasks had no effect. To examine the effect of the task description, two other groups of Black students were given the difficult task, were randomly assigned to a Black or White experimenter, and were told that the task was a test of intelligence. The researchers found that students who had the intelligence test task description scored lower when the task was presented

with a White experimenter, and when the Black experimenter presented the materials, the scores were slightly elevated.

Katz, Roberts, and Robinson (1965) also believed their results supported the Yerkes-Dodson Law (Yerkes & Dodson, 1908) based on three factors. One, an out-group experimenter produced more drive than an in-group experimenter; two, the intelligence test task description motivated students more than the motor-coordination description; and three, the effects of the task description and experimenter were additive. When the task was described as a motor skill, Katz et al. proposed that the Black experimenter did not motivate the students as much as the White experimenter. The task of intelligence was thought to arouse more motivation than the motor skill task description. When the task was introduced as a task of intelligence and there was a White experimenter, the situation may have no longer been motivating, but instead may have been overwhelming to participants. In comparison, when the experimenter was Black and the task was described as about intelligence, the participants may have been motivated at a level closer to ideal.

Research has also been conducted on prevention of ST related to gender using other task descriptions. In a series of studies, Quinn and Spencer (2001) examined whether ST affected women's ability to create problem-solving strategies. In the first study, 54 men and 54 women college students completed one of two different versions of the same test, (a) mathematical word problems or (b) the same mathematical problems in numerical form. The 2 (gender) x 2 (test type) ANOVA was statistically significant for the interaction effect. Contrast analyses indicated that both genders performed similarly on the numerical test, but women underperformed in comparison to men on the word problem test. Quinn and Spencer (2001) concluded that women had the skills to solve the mathematical problems, but that something interfered with their ability

to strategize and convert the word problems to numerical form. The interference was assumed to be due to ST.

In the second study, Quinn and Spencer (2001) examined how ST interfered with problem solving in a sample of 36 college students who scored between the 90th and 95th percentiles on the math portion of the SAT. The research design was a 2 (gender) x 2 (ST vs. low ST) between-subjects experiment. Participants were assigned randomly to one of two ST groups. The high ST group received the same type of introduction as in Study 1. In the low ST group, participants were told that the test to be administered would be gender fair. The ANOVA was not statistically significant for main effects, but the interaction effect was interpreted to be marginally statistically significant ($p = .07$). Contrast analyses of the interaction effect indicated that men performed better than women in the high ST condition, but no gender differences were found in the low ST condition.

Quinn and Spencer (2001) concluded that when the test was presented differently the meaning of the situation changed. Women in the low ST condition did not seem concerned about being judged (were not distracted by this concern) in comparison to the women in high ST condition, even though the difficulty of the test was the same. Although explaining the task as non-diagnostic or nondiscriminatory is a promising method to reduce ST, sometimes it is not possible to reframe the meaning of a test. High stakes tests like the SAT and GRE would be difficult to reframe, because most students are familiar with the diagnostic aspects of the tests and the implication of their use. However, Good, Aronson, and Harder (2008) claim that ST can still be reduced through explaining that the test is fair even if the test is diagnostic.

Making self-affirmation. Instructing people to think about their positive characteristics may protect people from ST (Bowman, Wegmann, & Webber, 2013; Schimel, Arndt, Banko, & Cook, 2004). This method of prevention is based on Steele's (1988) theory of self-affirmation. One goal of the self is to protect itself from threats to adequacy. As a result, Steele contends that when a person experiences threat to the self, the effects may be reduced through affirming other valued aspects of the self. In essence, self-affirmation is a method of reducing the impact of a threat that is more adaptive than denying or ignoring the threat (Sherman & Cohen, 2006). For instance, students who identify as a good student may reduce the threat to self when having difficulty in language arts by affirming that their strengths in another academic area.

Steele (1975) examined the effect of self-affirmation on 60 women who lived in middle- to upper- income neighborhoods. Women were randomly assigned to one of four conditions based on caller's perception of the community, including the participant: (a) the relevant negative condition described the community as not helping others; (b) the relevant positive condition described as helpful; (c) the irrelevant negative condition described the community not supportive of safety; and (d) the control condition received no telephone call. Under the first three conditions, women were called on the telephone and told that the call was a preliminary effort that may lead to an eventual poll concerned with women's opinions. Then their community was described based on the respective condition, given sample questions, and asked whether they wanted a poll reflective of the sample questions. Two days later in a seemingly unrelated telephone call, women were asked to help with a community project to create a food co-op by listing everything in their kitchen to help the co-op determine wholesale buying. It was found that the relevant negative name condition prompted more helping than the control or the

positive condition. In addition, the irrelevant condition prompted as much compliance as the relevant condition.

Steele and Liu (1983) further tested whether the affirmation of a valued self-concept could reduce dissonance even if the valued self-concept was unrelated. Dissonance is a type of threat to the self that creates inconsistency in the way a person views the self (Festinger, 1957). To reduce dissonance after individuals have done something counter to their beliefs, they will change their views to closer resemble their actions (Steele & Liu, 1981; Steele, 1988). In essence, to reduce the threat, it was expected that one must cast the self in a more positive light. University students were selected as participants based on either high or low economic political value orientations and then were randomly assigned to a low or high choice group. Students in the high choice group were asked to write an essay supporting a tuition hike, because the researcher already had more than enough essays against the tuition hike. This condition was meant to create dissonance in the participants, because of the choice to write this essay was counter to their beliefs. In the low choice group, students were told it was a requirement for them to complete an essay supporting the tuition hike. After completing the essay, students were randomly assigned to take a survey on economic political value orientation and then asked to complete an attitude survey. The economic political value orientation survey was intended to be a method of self-affirmation for students who had a high economic political value orientation. Steele and Liu found that students who had a high economic political value orientation and were able to affirm themselves were less likely to change their views on the tuition hike in comparison to the students who had a choice of essay. In other words, students who felt they had a choice in writing the essay and were able to self-affirm experienced lower dissonance than students who also had a choice to write an essay but were not able to self-affirm. Steele and Liu (1983)

demonstrated that self-affirmation could reduce threats, such as dissonance, even when the threat was unrelated to the affirmation, as long as the self-affirmation was of value to the person.

In a series of studies, Schimel and colleagues (2004) examined self-affirmation as a prevention method of ST. In the first study, the focus was on reducing student's self-handicapping behavior using a self-affirmation technique. The sample consisted of 11 men and 38 women college students, who were told the study was examining cognitive processing, and were assigned to one of three self-affirmation conditions (intrinsic, extrinsic, or neutral). The intrinsic affirmation condition asked students to describe things they valued about themselves that were not connected to imposed standards or external rewards. Those in the extrinsic affirmation condition were to describe things they valued about themselves that were attached to external rewards of some kind. The neutral condition consisted of students describing their daily activities.

Told that they were part of a study that examined personality and its relation to cognitive processing, students were asked to complete sentence stems, which varied depending on the experimental condition, then to complete a measure on self-handicapping, and finally to complete mental math problems. A 3 (self-affirmation condition) X 2 (gender) ANOVA on self-handicapping behavior revealed a statistically significant main effect for the self-affirmation condition. Students in the intrinsic condition reported less self-handicapping than those in the extrinsic affirmation condition or neutral affirmation condition. No differences were found in reported self-handicapping between the extrinsic and neutral affirmation conditions. In terms of task accuracy, a 3 (self-affirmation condition) X 2 (gender) ANOVA revealed a statistically significant main effect for the self-affirmation condition; students in the intrinsic condition had

higher accuracy than those in the extrinsic condition and the neutral condition was not statistically significantly different from the extrinsic condition and the intrinsic condition.

For Study 2, Schimel et al. (2004) focused on only the extrinsic and intrinsic groups, and manipulated ST. The sample consisted of 46 women college students, who were given a test booklet in which the first page varied depending on the experimental condition: (a) title of “quantitative examination” for the threat condition and (b) title of “problem solving exercise” for the no threat condition. The same instructions were given to all participants. Participants completed the math task and then completed a questionnaire on self-esteem. A 2 (ST condition or No-ST condition) x 2 (extrinsic self-affirmation or intrinsic self-affirmation) ANOVA revealed an interaction between self-affirmation and the threat conditions. Participants who received extrinsic self-affirmation were not as accurate in the ST condition as those with extrinsic affirmation in no threat condition. Students who received intrinsic affirmation in the ST condition were more accurate than in those with the intrinsic affirmation and no threat condition. Additionally, students with an extrinsic affirmation in the no threat condition were more accurate than students with an intrinsic affirmation in the no threat condition. Schimel et al.’s (2004) findings were used to provide support for using intrinsic affirmation to prevent or reduce the effects of ST on women completing math tasks.

Martens et al. (2006) also found support for the use of self-affirmation to reduce the effects of ST. Seventy-seven women and 70 men college students were randomly assigned to one of two conditions (No-ST and ST). The task was described as reasoning problems in the no threat condition and as assessing math and reasoning abilities in the ST condition. Women assigned to the ST condition were randomly assigned to receive either the self-affirmation condition or not. In the self-affirmation condition, participants were asked to rank order

characteristics and then write an essay about the one characteristic most important to them. In all the other conditions, participants were asked to rank order the characteristics, and then were asked to write an essay about their ninth most important characteristic and its importance to other people. Planned analyses showed that women in the self-affirmed group performed better than women in the ST group without affirmation. Martens et al. (2006) concluded that women's performance in mathematics does not decrease when in danger of experiencing ST, when women are instructed to self-affirm.

Self-affirmation as a method of ST prevention has also been supported outside of the laboratory setting. Cohen, Garcia, Apfel, and Master (2006) tested a values-affirmation intervention based on Steele's (1988) theory of self-affirmation to prevent stereotype threat. In a two-wave field study, the sample consisted of 119 African Americans and 124 European American students from two successive years of seventh grade. Students were randomly assigned in the fall term to either the control or the affirmation condition. From a list of qualities, the control condition students were asked to indicate their least important quality and write why it was important to someone else. In contrast, the students in the affirmation condition were told to write about their most important qualities and why these were important to them.

A multiple regression was conducted using the predictors of race and condition, with the dependent variable the students' letter grade (grade metric A = 4 and F = 0) in the class. In the first and second waves, a race x treatment interaction effect was found. African American students in the affirmation condition had higher grades than African American students in the control condition. No treatment effect was found for the European American students. In addition, a two-year follow-up study was conducted. Cohen, Garcia, Purdie-Vaughns, Apfel, and Brzustoski (2009) found that, on average over the two years, African American students' grade

points raised by 0.24 grade points. Among the students who had been affirmed, it was also found that the number of African American students identified as at-risk and enrolled in remediation had decreased.

Self-affirmation appears to be a promising method of ST prevention. However, specific criteria must be met in order to prevent ST. Self-affirmation appears to be more effective when the affirmation is about something of value to the person (Steele & Liu, 1983), and only if the valued concept is intrinsically motivating to the person (Schimel et al., 2004). Thus, to activate the prevention response to ST, students need to be aware of what is intrinsically motivating to them in order to self-affirm. In addition, teachers need to be aware that self-affirmation is an effective method of ST prevention and be knowledgeable about their students to help them find appropriate characteristics about themselves to self-affirm to prevent ST.

Providing a role model. Another way of reducing the effects of ST is to provide a role model to the students who demonstrate strength in a domain. The premise behind this method of prevention is based on two social psychology concepts: diffusion of responsibility and the identification contrast model. Diffusion of responsibility (Darley & Latane, 1968) is when a person is less likely to feel responsible in a situation when other people may also be held accountable. Typically, level of responsibility decreases as the number of people increases (Darley & Latane, 1968). McIntyre and colleagues (2003) contend that telling students about successful in-group people would create a climate of diffused responsibility. In essence, the successful in-group members would serve as role models and relieve the pressure of students feeling responsible to achieve to overcome the designated stereotype of their group. The role models would diffuse responsibility, thus leading to the prevention of ST.

The identification contrast model (Buunk & Ybema, 1997) has also been used to explain how to reduce ST. The premise of this model is that people are motivated to relate to others who are doing better than they are. People tend to focus on shared attributes with others, and thus see others' accomplishments as attainable for themselves (Buunk & Ybema, 1997). In essence, people may benefit from hearing about successful individuals who are similar to them (Ybema & Buunk, 1997). Marx, Ko, and Friedman (2009) indicated that telling students about a role model who is perceived as in-group and successful would help students believe that they too could defy the stereotypes like the role model. Thus, students viewing themselves as similar to the role model may believe that they too could achieve success; in turn, ST is prevented.

Studies (Marx, Ko, & Friedman, 2009; McIntyre, Paulson, & Lord, 2003) have provided support for using role models to prevent ST. In a series of experiments, McIntyre and colleagues (2003) examined whether ST of gender could be reduced. In the first experiment, college students (116 women and 46 men) were randomly assigned to one of two conditions (control and successful women). In both conditions, a woman researcher told participants that some research showed that men outperformed women in math to induce the gendered math stereotype. In the successful women condition, the packet instructions contained a section noting that in future experiments only women will be invited to return, because their data were more valid and reliable. In the control condition, no announcement was given about participant selection in future studies. Students then had 20 minutes to complete 34 difficult items from the GRE. Then students completed a measure on perceptions of their performance and reported their SAT and GRE scores. A planned comparison was completed on the 2 (gender) by 2 (condition) between subjects design and the findings revealed that women in the control group performed worse than

those in the other groups (men in the control condition, men in the successful women condition, and women in the successful women condition).

In the second experiment, the sample consisted of 74 women and 32 men college students (McIntyre et al., 2003). In this experiment, a man researcher informed students that men were better at math than women before the students were asked to read and critique four essays. Students were assigned to one of two conditions (control and successful women). In the successful women condition, students had to read four fictitious essays about four successful women and their careers. Students in the successful corporation condition (control) read about four successful corporations. Then the packets were collected and a woman entered the room and had the students complete a math task and questions that were presented as a separate experiment from the essay critique. Findings of planned comparisons on the 2 (sex) by 2 (condition) between subjects design revealed that women in the control group performed worse than the other three groups (men in the control condition, men in the successful women condition, and women in the successful women condition). McIntyre et al. (2003) concluded that general and specific reminders of women's achievements seem to prevent the negative effects of ST.

Marx, Ko, and Friedman (2009) also examined whether the effect of a single role model (Barack Obama) through positive media attention helped to prevent effects of ST. Participants were drawn from a nationwide wide sample of Black ($n = 84$) and White ($n = 388$) adults, who were matched on language proficiency and age. There were four points of data collection and each point included about a fourth of the participants. At each data point, an ANCOVA that controlled for educational level was conducted. Data collection periods 2 (after President Obama was nominated at the Democratic National Convention) and 4 (after President Obama was

elected into the Office of President) were selected when Barack Obama's success was more salient. Data collection periods 1 (prior to the convention) and 3 (the mid-point between the convention and the election) were selected for when Barack Obama's successes were less salient. During data collections, participants were asked to report their race and told they were taking a 15-minute assessment (GRE Verbal section) that examined intellectual strengths and weaknesses.

For the first data collection period (prior to convention), White students were found to have higher scores than Black students. In the second data collection period (Obama's nomination), Marx and colleagues (2009) expected that only the Black participants who watched Barack Obama's acceptance of the nomination speech would be positively affected. White participants scored higher than Black participants on the verbal section of the GRE. An interaction effect was found between race and whether or not the convention speech was watched. White participants who did not watch Obama's speech scored higher on the GRE Verbal section than Black participants who did not watch the speech. However, Black and White participants who watched the speech scores did not differ from each on GRE Verbal scores.

For the third time of data collection (post-convention, pre-election), when Obama's success was less concrete, White participants performed better than Black participants. At the fourth data collection, which was held after the election, Marx et al. (2009) found no effect of race on performance on the GRE Verbal section. It was concluded that the research provides support that role models may provide a protective factor against the effects of negative stereotypes.

Providing a role model for students could be a useful strategy in the prevention of ST. However, certain conditions must be met to ensure the effectiveness of role models against ST. The identification contrast model requires that a person must be considered to be an in-group

member (Lockwood, 2006) who is successful (Marx & Roman, 2002). Thus, teachers must be aware whether students perceive the in-group member as a role model and as successful in order to use this method of ST prevention.

Switching identities. Some studies (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004; Rydell, McConnell, & Beilock, 2009) have shown that reducing the salience of the identity that is threatened may reduce ST. The premise behind switching focus to another identity is related to social identity theory (Tajfel & Turner, 1986). The basis of social identity theory is that people are motivated to see themselves positively, which is accomplished through identifying with social groups. One way to protect the self when a group's stereotypes threaten a positive view of the self is to leave the group temporarily or permanently by identifying more strongly with another social group (Tajfel & Turner, 1986). In this way, the negative stereotypes will no longer apply, and ST will be prevented, because the self identifies with another social group who is associated with more positive attributes or stereotypes.

In one of the first studies to examine this concept, Ambady et al. (2004) examined whether people who engaged in personal disclosure (focusing on the self rather than on a group identity) were less likely to be negatively affected by ST. In the first experiment, 40 women college students were randomly assigned to one of four conditions based on a 2 (prime: gender or neutral) by 2 (questionnaire type: individuated or non-individuated) between subjects design. Participants completed a task on the computer and were asked to indicate where on the screen a word appears. Women primed for gender were given words like jewelry, lady, and lipstick. Women not primed for gender were given words like water, table, and stapler. Women also were assigned to receive one of two questionnaires: the individuated questionnaires asked the participants about things like their favorite book and favorite hobbies; and the non-individuated

questionnaires contained questions about lions. After completing the questionnaires, students completed a challenging math task. Based on a two-way ANOVA, women who were primed for gender and in the individuated condition performed better on math problems than those who were primed for gender and in the non-individuated condition. Women who were primed for gender and in the individuated condition did about as well as the women who were not primed and not in individuated condition.

The second experiment, a replication of the first study with 39 women Caucasian students, yielded similar results. Women who were primed for gender and were individuated performed better than women who were primed and not individuated. No differences were found between both unprimed conditions (individuated and non-individuated) and the gender primed and individuated condition. Ambady and colleagues (2004) concluded that individuation appears to play a role in the prevention of the effects of ST.

Rydell, McConnell, and Beilock (2009) also examined the effects of gendered STs by giving women college students' cues about their identity. However, the focus was on whether activating a favorable group identity in the presence of a negatively stereotyped group identity would prevent ST. Activating a favorable group identity was believed to prevent ST, because one of the assumptions of social identity theory is that people are motivated to view themselves positively (Tajfel & Turner, 1986). When given two identities, to protect the self, a person will choose to identify with the more favorable group identity. Rydell and colleagues (2009) assigned 122 women undergraduate students to one of four groups. One group was given no information about math performance stereotype (control condition); another group was told that the purpose of the study was to examine why women are generally worse at math than men. The third group was told the purpose of the study was to examine why college students are generally better at

math than non-college students, and the last group was told that the purpose of the research was to find out why men are better than women at math and why college students are better at math than non-college students. The students then completed a math task. Based on a 2 (gender stereotype) by 2 (college student stereotype) ANOVA on the number of math problems correctly answered, an interaction between gender stereotype and college student stereotype was found. Women students who were presented with gendered stereotype information and received the college stereotyped information performed better on the math task than did students who received gender stereotype information, but did not receive the college student stereotyped information. No difference was found in women students' math performance in the groups not presented with gender stereotype information (women presented college student stereotype and women not presented with college student stereotype). Rydell et al. concluded that presenting an alternative positive stereotype in the presence of a negative stereotype could prevent the effects of ST.

Switching focus to another identity that is not threatened by ST could be a useful method of ST prevention for teachers. Helping students focus on their individual strengths or identities may be less practical in a classroom setting, because it could be difficult and more time consuming to infuse into class time for each student and may only be possible on a one-on-one basis. However, teachers could reinforce alternative identities that are not negatively stereotyped (Rydell et al., 2009). Switching focus to another identity may be particularly useful for three reasons: (a) It may still prevent ST even when negatively stereotyped identity is activated, (b) it can be infused into everyday lessons, and (c) it can be used as a class-wide method of prevention, by using broad identities that everyone in class can identify with (e.g., "students").

Viewing the brain as a muscle. Research has supported that ST may be prevented when people believe the brain is a muscle and intelligence can grow (Aronson, Fried, & Good, 2002). The basis for this method of ST prevention is related to how students view intelligence. Dweck (1986, 1999) conceptualized that students' beliefs about intelligence can have an impact on their goals, reaction to experiencing difficulty, and academic achievement. Students can hold either an entity or an incremental view of intelligence. Students who have an entity view of intelligence tend to be interested in performance and doing well; thus, they are more interested in tasks that will demonstrate their ability to perform well. When something is difficult, these students tend to become uninterested or incapacitated. In contrast, students who have an incremental view of ability tend to be interested in improving their skills and learning new concepts. When something is difficult, these students try harder and interest is heightened (Dweck, 1986, 1999). Dweck (1999) believes that students with an incremental view of intelligence also tend to feel less anxiety when they experience difficulty, because their goals are related to learning instead of performance.

Aronson and colleagues (2002) examined the effectiveness of encouraging students to view intelligence incrementally as a prevention method for ST. Participants were 79 African American and Caucasian undergraduate students (men and women) who were randomly assigned to one of three writing groups (non-pen pal, malleable pen pal, and control pen pal); they were asked to write letters to help middle school students who were at-risk. People in the malleable pen pal condition were asked to write letters that advised students to work hard and that intelligence is expandable through mental work. In the control pen pal condition, students were asked to respond with encouraging letters and to also explain the different kinds of intelligence.

In the control group, the students were not asked to write a letter. After writing the letters the students completed measures on their identification and enjoyment of academics.

The design of the study was 2 (race) by 3 (condition) between subjects design. An ANCOVA with previous SAT scores as the covariate was conducted. Main effects for race and condition were statistically significant, but not for the interaction. Post hoc analyses indicated that African Americans were more likely to view intelligence as malleable in comparison to their Caucasian peers. In addition, students in the malleable letter group were more likely to view intelligence as malleable in comparison to the pen pal control and no letter control conditions. The ANCOVA on school enjoyment was statistically significant for race by condition interaction. African American students indicated enjoying school more when writing a letter in the malleable pen pal condition in comparison to those in the control pen pal condition. African American students were not found to differ in enjoyment of school between the control, no letter, and the control letter group. White participants reported more enjoyment in school in the non-pen pal control in comparison to the pen pal control.

Aronson et al. (2002) also examined the effect of writing conditions on academic performance using actual grades from transcripts from the following semester. Based on an ANCOVA, with SAT scores as the covariate, main effects for race and condition were found. White participants had higher grades than the African American participants. However, African Americans in the malleable pen pal condition obtained higher grades in comparison to African American in either of the other conditions. In contrast, no differences were found in grades between the conditions for White participants.

In another study, Good and colleagues (2003) examined the effect of incremental intelligence with 138 seventh grade boys and girls in a low-income rural school district. Eighty

percent of the participants were reported as being either Black or Hispanic; however no breakdown of the distribution of ethnicities was provided. The students were mentored by college students, who were randomly assigned to students and communicated with the students in person and through email throughout the year. In the incremental condition, the students learned how intelligence could grow over time; in the attribution condition the students learned how students in seventh grade at first will have difficulty, but will gain experience and improve. In the combination condition, students learned how intelligence could change and about the initial difficulty of seventh grade; whereas in the control condition known as the antidrug condition the students learned about the negative effects of drugs.

To assess the effectiveness of the intervention, Good and colleagues (2003) used the Texas Assessment of Academic Skills (Texas Education Agency, 2003). Based on 2 (gender) by 4 (experimental condition) ANOVA, an interaction effect on math performance was found. Planned comparisons indicated that boys in the control condition performed better than girls. Good and colleagues (2003) reported that there were large improvements in the incremental, attribution, and combined conditions for girls in comparison to those in the control condition.

When reading scores were used as the outcome measure, the ANOVA showed a main effect for condition (Good et al., 2003). Incremental condition students and attribution students scored higher on the reading portion of the test than students in the control condition (Good et al., 2003). No difference was found between those in the combined condition and the other conditions. Good et al. concluded that interventions including an incremental view of intelligence were promising in the prevention of the effects of ST.

Encouraging students to use an incremental view of ability could be a useful way for teachers to prevent students from experiencing ST. However, this prevention method so far has

only been shown to be effective only through repeated exposures (Good et al., 2003). In a school setting, this prevention method requires a teacher to expose students to the idea of incremental intelligence over time in order for students to internalize the view themselves. No research has been conducted to support that this method of prevention can be introduced once right before a challenging task and still be effective. Thus, teachers will need to infuse this concept into their teaching as often as possible.

Providing external attributions. Research supports that providing students with explanations for their anxiety prevents the effects of ST (Johns, Inzlicht, & Schmader, 2008; Johns, Schmader, & Martens, 2005). This premise is based on appraisal-based models of stress (Lazarus, 1991). When people are in an environment that is confusing or unclear and there is a possible threat to the self, they are motivated to control or suppress their feelings and thoughts (Johns et al., 2008). It is thought that this method of coping can expend mental resources needed for the task at hand and may result in decreased performance (Schmader et al., 2008). Thus, it is believed that when an explanation of the situational difficulty is given, such as an explanation for anxiety, those feelings no longer need to be repressed and thus preventing the process that enables ST to decrease performance.

Johns et al. (2008) examined the impact of telling students about anxiety in math in an effort to prevent ST. Participants, 58 Caucasian women college students, were initially told that the purpose of the research was to examine tutoring practices. A confederate was always assigned as the tutor, a woman in the control condition and a man in the ST condition. Once the confederate left the room, participants looked over the sample math problems. In the control condition, the math problems were called problem-solving tasks. For the ST condition, the task was described as a math test. In the stereotype plus anxiety reappraisal condition, participants

were informed that research had shown that anxiety did not hurt math performance and may even help it. Then the students completed a math task. Students in the ST condition answered fewer questions correctly than in those in the control condition or in the anxiety reappraisal condition. The control condition and the anxiety reappraisal condition did not differ on math performance. It was concluded that giving participants a way to deal with emotions through reappraisal of anxiety could be an effective method of preventing the effects of ST.

Using external attribution for difficulty as a prevention method for ST, Johns et al. (2005) taught participants about ST and its possible effects before taking a math test. The sample contained 75 women and 42 men White college students, who were randomly assigned to one of three conditions. Students in the math test condition were told that they would be taking a test to obtain information on gender differences. Students in the problem solving condition were told that they would be solving problems to obtain information about general cognitive processes. Students in the teaching-intervention condition received the same explanation about the math test condition, along with an additional explanation of what ST was; they were also told that the stereotype has nothing to do with their ability to do well on the test. Students were given 20 minutes to complete a math task. The researchers used a 2 (gender) by 3 (condition) between subjects design. ANCOVA, with previous SAT score as the covariate, was statistically significant for gender as well as gender x condition on math accuracy. In the problem solving condition, there were no differences between men and women, but in the math test condition, men were more accurate on math than women. There was no difference in math performance between women in the teaching intervention condition and men in the teaching intervention condition and women in the problem solving condition. Johns et al. (2005) concluded that teaching students about ST could protect against the effects of ST.

Providing external attributions for ST may be a useful strategy for teachers to use in order to prevent ST. However, teachers should be cognizant of the age of the students. The discussion of concepts like anxiety and ST may not be effective, because younger students may not be developmentally ready for such sophisticated concepts. The studies for this prevention method only used college students to test its effectiveness (Johns et al., 2005, 2008), so this method's effectiveness with younger students is unknown.

Summary. Six prevention strategies were reviewed that have been supported in the literature for reducing or preventing ST. One, explaining a task differently has been shown to prevent the effects of ST. For example, describing a task as gender fair reduced the effect of ST in women in comparison to women who were told no information about gender (Spencer, Steele, & Quinn, 1999). Changing the description of the test from diagnostic to non-diagnostic of ability was also found to reduce the effects of ST in African American samples (Steele & Aronson, 1995).

Two, self-affirmation may be an effective method of ST prevention. Schimel and colleagues (2004) found that asking people to think about their positive qualities could be a protective factor from ST. A third prevention is to provide students with explanations for their anxiety. Johns, Inzlicht, and Schmader (2008) found that when women were told that anxiety at times could improve one's math performances, their math performance did improve. Thus, giving an explanation about being anxious worked as a protective factor against ST. Additionally, Johns, Schmader, and Martens (2005) found that teaching participants about ST and its possible effects could also prevent the effects of ST.

Four, using a role model has been found to be an effective method to prevent ST. McIntyre et al. (2003) found that general and specific reminders of women's achievements may

prevent the negative effects of ST. Marx, Ko, and Friedman (2009) also found support that role models in current events could prevent ST with an African American population. Five, switching identities has been another way to prevent the effects of ST. Ambady, Paik, Steele, Owen-Smith, and Mitchell (2004) found that when women engaged in thinking about their individuality, they were less likely to be negatively affected by ST. Finally, how intelligence is viewed seems to have an effect on preventing ST—emphasizing the plasticity of intelligence and its growth overtime. African American students who were asked to write about intelligence as being plastic obtained higher grades than African American students in the control condition and in the no write condition (Aronson et al., 2002).

Wise Interventions

There is an emergent class of interventions called wise interventions. Walton labeled them “wise to specific underlying psychological processes that contribute to social problems or prevent people from flourishing” (p. 73). The focal point of a wise intervention is to first identify the psychological process at hand. These interventions are based on a lengthy history of research on psychological processes, but the characteristics of these interventions make them new. One, the interventions are intended to be psychologically precise. Two, they are expected to be brief. Three, the use of these interventions is to change self-reinforcing processes that occur over time. And, four, they are designed to improve individuals’ long-term functioning under various circumstances (Walton, 2014). Walton (2014) provides an in-depth review of the wise interventions.

Some of the ST prevention methods contained in the literature review are wise interventions. For instance, the incremental intelligence approach (Aronson et al, 2002; Good et al, 2003) is a wise intervention, because an incremental view of intelligence helps students

approach challenging course content as learning opportunities rather than of evidence of a fixed intelligence (Walton, 2014). Another example would be the approach to providing an explanation for the anxiety, because an explanation for the anxiety allows a person to free up mental resources that may have otherwise been spent repressing the anxious feelings (Johns et al., 2008).

Wise interventions are thought to disrupt cycles of threat (Walton, 2014). However, timing is important, because interventions have been found to be less effective when the negative cycles have already been established (Cook, Purdie-Vaughns, Garcia, & Cohen, 2012). In terms of ST, a prevention or wise intervention method may be more effective in the elementary grades than at other ages, because by the age of 5, most children are aware of racial and ethnic stereotypes (Aronson, 2009) and are susceptible to internalizing stereotypes and prejudices (Martines, 2008). In addition, early intervention is imperative to prevent chronic ST, which can cause a student to disidentify with a field related to the ST (Griffin, 2002; Steele, 1997). Students have been found to experience ST as young as 6-10 years of age (McKown & Weinstein, 2003). It is believed that when ST is experienced by young students, preschool and elementary, by the time they reach high school, they may have already disidentified from academics due to chronic ST (Jordan & Lovett, 2005). Thus, it is plausible that many college students may have already disidentified from academics or certain fields of study before college due to chronic ST. At the same time, college students have benefitted from the ST or other psychological research. A majority of the research on ST and prevention has been conducted on college aged students (e.g., Ambady et al., 2004; Martens et al., 2006; Quinn & Spencer, 2001; Rydell et al., 2009). Steele (1997) used the term “wise schooling” to emphasize the need for school personnel to become sensitive to the experiences of children from diverse backgrounds in order to mitigate the effect

of stereotype threat. Thus, it is important to study ST prevention methods or wise interventions early in students' academic careers to prevent chronic ST and its effects.

Teacher Change

To achieve the effects of wise schooling, it is critical to develop interventions that will involve teachers. Whether interventions are conducted across a school system or within a specific classroom, implementations of any school interventions rely on the buy-in of teachers and their motivation and willingness to follow through. Any school interventions also require that teachers have received the necessary knowledge to be effective in implementing them.

Guskey (2002a) proposed a model of teacher change related to professional development. The premise behind his model of teacher change was based on James' (1890) conceptualization of the connection between behavior and emotion change. For example, if a person sees a threatening animal and runs, and as a result of such behavior, the person would attribute that behavior to being afraid. In this way, teachers receive a professional development program, then these teachers change their classroom practices based on the professional development. If positive student outcomes are perceived by the teachers, then that experience will change teacher beliefs and attitudes. Guskey (2002a) stated that observable positive outcomes are the catalyst to any enduring instructional change. There has been little research that has examined this model of teacher change (e.g., Hodges et al., 2017), because the majority of the research related to professional development has been dedicated to examining the effectiveness of the professional development (e.g., Davis, Preston, & Sahin, 2009; Guskey, 2002b; Owston, Sinclair, & Wideman, 2008) rather than the mechanisms that promote change in teachers (e.g., Hodges et al., 2017).

Hodges et al. (2017) applied and examined Guskey's (2002a) Model of Teacher Change as a framework to improve teacher's instruction of physical education. Nine physical education teachers took part in semi-structured interviews and five teachers were randomly assigned to a professional development curriculum to improve teacher's health-fitness instruction, which involved specific training sessions, YouTube videos that modeled lessons, and teaching materials. The teachers were interviewed two additional times, one week after the teacher started using the professional development materials and one month after the project was completed.

Hodges et al. (2017) identified three themes that appeared to support the use of Guskey's model (2002a). The first theme was that teachers acknowledged the importance of health-related fitness instruction, but reported insufficient time or resources to cover the material. This theme seems to represent the first step of the model—teachers receiving professional development training. The second theme was related to teachers making changes after the training, which is similar to second step of Guskey's model—teachers' incorporating the professional development into their teaching practice. The third theme identified was that professional development was deemed beneficial based on observed student outcomes. This step also mirrored Guskey's model—teachers' observations of students' benefit to the intervention. Hodges et al. also reported support for the final step of Guskey's model: Teacher change of thoughts and attitudes based on their intent on continuing to use what was learned through the professional development training. This finding provides promising support for Guskey's model, but additional research on this model of change is needed. Thus, educating teachers about prevention methods to mitigate ST effects may be the first step in promoting lasting benefits to students who are experiencing ST about academics.

The Current Study

Research on ST is vast, but little has been conducted in real-world settings, such as the classroom (Good et al., 2003; Keller, 2007; Stricker & Ward, 2004). Furthermore, methods to prevent ST have been supported through several studies, but none have focused on teachers' ability to use various ST prevention strategies. In general, no studies have been found that focused on teachers' knowledge of ST and the impact of such knowledge in preventing ST. The purpose of this study was an initial step toward investigating pre-service teachers' learning about ST and reducing its impact on students. Particularly, the aim of the study was to gain knowledge on how pre-service teachers would respond to a short lesson on ST, demonstrate knowledge about ST, and then select an appropriate method of ST prevention. The following research questions guided this study:

1. When exposed to the ST video module, did pre-service teachers gain more knowledge about ST than pre-service teachers who were not exposed to the ST video module?
2. When exposed to the ST video module, were pre-service teachers more likely to select appropriate prevention strategies between an ST scenario and a No-ST scenario in comparison to pre-service teachers who were not exposed to the ST video module?

The research questions above led to the formulation of the following hypotheses:

Hypothesis 1: Pre-service teachers exposed to the ST video module were expected to have a higher score on ST items than pre-service teachers who had not been exposed to the ST video module.

Hypothesis 2: Pre-service teachers exposed to the ST video module would rate on average the ST effective strategies as more likely to use for the ST scenario in comparison to pre-service teachers who were not exposed to the ST video module.

METHOD

This section will cover the methodology for two pilot studies and the current study. Unless specified, the methodology will be described for each pilot and the current study.

Pilot 1

The purpose of the first pilot was to ensure that the ST video module was accessible to the participants. Furthermore, Pilot 1 was conducted to obtain feedback about the construct validity of the items, scenarios, and ST video module.

Research Design

The research design of the experiment was a 2 x 2 between subjects design, which involved two manipulations. The first manipulation was the intervention: Exposure to an ST video module or not. The second manipulation was the type of hypothetical scenario given: ST or not.

ST video module. Exposure to ST knowledge was manipulated. In the control condition, no information was provided about ST. In the treatment condition, a module on ST knowledge was provided. The treatment module was a video providing knowledge about ST and lasted 6 minutes and 48 seconds. The video introduced the concept of ST and gave an overview of the situation in which ST may be experienced (students' investment in an area of performance, awareness of a stereotype, and acceptance of the stereotype as true about their group and themselves through their actions). Then a short explanation of the importance of preventing ST was given. These two portions of the video lasted for about 3 minutes and 10 seconds. Then 2 minutes and 35 seconds of the video highlighted six strategies that teachers could use to prevent students from experiencing ST. Then, the final 1 minute and 3 seconds reviewed the contents of the video.

Scenarios. The second manipulation involved the type of hypothetical scenario presented to participants: ST scenario and No-ST scenario. The purpose of this manipulation was to test whether the exposure to ST knowledge was effective in teaching pre-service teachers about the concept and its application. The scenarios were brief and depicted a student who was having difficulties in school. In the ST scenario, the student was struggling in mathematics due to experiencing ST, evidenced by the student's investment in the area of performance, the awareness of the relevant stereotype, and the awareness of confirming the stereotype (Steele & Aronson, 1995). The ST scenario is presented below:

Sophia¹ is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is worried about the math unit test. She stated, "Tests in math are hard. I have to do good or else I can't go on the science field trip and then if I can't go to that I'll never be a scientist." The field trip was too expensive for some students' families, but, fortunately, your school offered field trip funds to students who were doing well in all their subjects, but could not afford to pay for field trips. Sophia's parents shared with you earlier that week that she would be unable to attend the field trip unless the school could fund her way. In the past, you have noticed that Sophia is a bad test taker in math. She works hard in your class and does a nice job on her assignments, but when you grade Sophia's math tests they do not seem to accurately show what she knows. Sophia has strong problem solving skills and is fluent with simple addition, but on tests, addition seems to take her a much longer time and she has difficulty determining appropriate strategies to use even though she does

¹ An online search revealed that the most popular name for female babies born in 2011-2013 in the United States was Sophia (Social Security, 2014).

not have this difficulty during class assignments. On her most recent math test, she appeared to not have enough time to complete the test, because several questions towards the end were left blank. When you asked Sophia about her most recent math test, she stated “I tried really hard, but my brother says ‘Girls can’t do math.’ I don't think I got a good grade and now everyone will think girls can’t do math.”

In the scenario, investment in the area of performance is evidenced by Sophia’s comment about her desire to be a scientist. Awareness of a gender stereotype is alluded to when Sophia stated that her brother says, “girls can’t do math.” Her fear of the stereotype is confirmed in that (a) Sophia does not believe she has done well on the test, and (b) that everyone will think girls cannot do mathematics.

In contrast, the No-ST scenario involved a student who lacked motivation in mathematics. A lack of motivation in a subject matter was selected, because Steele (1997) stated that lowered test performance differences between ST and control groups were due to the pressure of ST and not due to motivation. The No-ST scenario is presented below:

Sophia is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is not concerned about the math unit test. She stated, “I don’t need to do math. I’m going to be famous, and famous people don’t go on science field trips.” Your school offers field trip funds to students who cannot afford to pay for the field trip and are doing well in all their subjects. Sophia’s parents shared with you earlier that week that she would be unable to attend the field trip unless the school could fund her way. Unfortunately, Sophia may not be

able to get the funds needed to attend the field trip due to her poor marks in math and science.

In the past, you have noticed that Sophia is not completing the assignments. Even when she is given extra time in class to complete her work, you need to remind her often to stay on task and to stop talking with her friends. When you grade Sophia's math homework and her tests, few problems are completed. Sometimes she quickly and correctly completes tests, and other times she will turn in a page filled with doodles. Whenever her test is filled with doodles, you always give her a new test and make her complete it. On her most recent math test, she quickly completed her test and when you checked it most of the problems were correct and all were attempted. When you asked Sophia about her most recent math test, she stated, "I wanted to draw. You would have made me do the test again if I drew on the test, so I just did the test fast so I could draw."

To ensure that this scenario did not depict ST, several changes were made. In this scenario, Sophia is not invested in math, because she expresses that she does not need math. Sophia does not appear to be aware of a ST, because information about related stereotypes are not mentioned and are purposely excluded. She also does not appear to be concerned or mention any concern about confirming any stereotypes.

Participants

Pilot 1 sample had six respondents, who identified as Caucasian women and employed by a suburban public school district in Maryland. Their ages ranged from 29 to 35 ($M = 31.67$; $SD = 2.34$). All reported having a family income of \$45,000 or more. All participants were teachers

who taught grades from kindergarten to fifth grade. Three teachers indicated they had “6 to 10 years” of teaching experience and three teachers reported having “11-15 years” of experience.

Measures

Four sets of information were collected: (a) demographic information, (b) knowledge about the concept of ST and other concepts relevant to teaching, (c) knowledge to identify a ST scenario and identify appropriate interventions methods, and (d) opinions on the video module and items presented.

Demographic information. Demographic variables about participants were collected. These variables were race/ethnicity, gender, age, class standing or lack thereof, socioeconomic status, GPA, actual grade level(s) taught, and number of multicultural conferences/classes/workshops attended.

Test of Application. The primary investigator created a measure called the Test of Application (TOA) for the study. It was used to assess participants’ ability to select appropriate strategies in response to a scenario. The TOA was composed of 11 possible prevention methods that could be selected to use with the student in the hypothetical scenario. These questions were presented after the hypothetical scenario. Participants were asked the following question: “As Sophia’s teacher what would you put in place to help Sophia achieve to what you think she is capable of in math?” The teachers then were asked to respond to the options for possible ST prevention (or lack of), rated on a scale from 1 (*Unlikely to Do*) to 5 (*Likely to Do*). Three methods of prevention were from the ST video module, three prevention methods were generalized from the ST condition, four methods of prevention were used not shown to be efficacious ways of intervening on ST, and one option was that no action would be taken. An example of a prevention method from the ST video module is, “remind the students of a positive

identity before math activities and tests” (switching the focus from the threatened identity; Ambady et al., 2004). An example of a prevention method that was generalized from the ST video module is, “Have your students take turns giving each other praise.” This item originates from the concept that self-affirmation is an effective prevention of ST (Taylor and Walton, 2011). A sample item of an ineffective ST prevention option is, “Create math activities that are related to her interests.” If Sophia is experiencing ST, she is already invested in mathematics (Steele & Aronson, 1995); thus this strategy would not be effective. A higher score indicated that a teacher was more likely to use that method of prevention when working with a student who presented with difficulties like those presented in the scenario.

Test of Knowledge. The primary investigator created a measure called the Test of Knowledge (TOK) for the study. The 24 TOK items were used to assess three knowledge areas (classroom organization, multicultural teaching, and ST). Eight questions tapped the content of the ST video module. ST questions were based on ST literature (e.g., Aronson, Fried, & Good, 2002; Steele, 1997; Steele & Aronson, 1995). An example of an ST item is, “Stereotype threat only affects performance in areas related to school.” The additional items were about classroom organization (8 items) and multicultural teaching (8 items). The classroom organization questions were intended to be distractor items and were based on material from McLeod, Fisher, and Hoover (2003), and Evertson (1989). An example of a classroom organization item is, “When organizing a classroom, it is important for teachers to identify the best location(s) to give instruction.” The multicultural teaching questions were based on material from Wilson (1997). An example of a multicultural item is, “It is not necessary for teachers to see the cultural background of their students in order to be effective as a teacher.” Respondents were asked to indicate whether each item was true or false. Thus, the correct options were rated as 1, and

incorrect options as 0. Total score was based on summing across all the ST items in that particular domain. Higher scores for the total ST scores were considered to reflect a beginning knowledge base about ST.

Procedures

The Pilot 1 sample was recruited using the snowball method, either through email or the social media platform Facebook. Participants were told that the purpose of the study was to examine the effectiveness of training videos to improve teacher effectiveness. To increase participation, four \$25 gift cards were offered through a raffle. The study was conducted through the online survey services of Qualtrics (2013). Those teachers who agreed to participate in the study received a web link via email or through a message on Facebook to access the study. After reading an implied consent form, the teachers answered demographic questions and then were randomly assigned to watch the ST video module or not. The video module was 6 minutes and 48 seconds in duration. Then, teachers answered TOK. All three sets, classroom organization, multicultural teaching, and ST, of items were randomly presented directly after the first manipulation. Then participants were randomly assigned to one of the two hypothetical scenarios (ST or No-ST). After reading the scenario, participants answered the TOA. As part of the pilot process, the participants were guided through the survey again, but this time the teachers were asked to answer open- and closed-ended questions about the content and clarity of the video module, scenarios, and other items.

Once participants had completed giving their feedback and to maintain their anonymity, participants were automatically redirected to a separate web link to register for the raffle. The raffle for the gift cards was conducted when data collection concluded and the winners were notified.

Results

The feedback from the participants was only examined qualitatively, due to the sample size (6). In regard to the accessibility of the ST video, both participants who viewed the video module reported that “the video was easy to access.” The other goal was to get feedback from the participants of Pilot 1 to improve the clarity of the items, scenarios, and video module. It was found that the majority of the items were reported to be clear. However, one item required alteration: a ST question in the TOK. This item was revised so it no longer contained a typographical error. The item was changed from “Someone needs to explicitly teach stereotype in order for the student to experience stereotype threat” to “Someone needs to explicitly teach a stereotype in order for the student to experience stereotype threat.” Another change made was in the instructions given to the participants. Two participants commented that the TOK sections of the survey were confusing, because instructions for each section were missing. Thus, each section was revised to include instructions.

Other Revisions to Research Design

Three additional changes were made to minimize participants from knowing the intent of the study before and during the study. One, the implied consent form contained information about multicultural teaching and ST, so this type of information was removed from the form to minimize its influence on the participants’ answers. Two, the order of the measures was also changed. The demographic section was moved from the first section to after the scenarios, because the question about the number of multicultural conferences might also prime the participants about the purpose of the study. The multicultural teaching questions were moved from after the video module to after the scenario, again to reduce priming participants to give more multiculturally sensitive answers.

Pilot 2

The purpose of the second pilot was to double check the clarity of the items, scenarios, and the newly added instructions for each section. Additionally, more detailed feedback was requested on the content, length, and quality of the video module.

Research Design

The research design for Pilot 2 was the same as in Pilot 1, a 2x2 between subjects experimental design. Furthermore, both manipulations, module and scenario, remained the same in this pilot as it was previously.

Participants

Pilot 2 consisted of 15 respondents who participated in the study. Of the sample, 60% of the sample identified as female, followed by 20% males, and 20% who indicated no gender. In terms of ethnic background, 93.33% of the sample identified as Caucasian. The remaining person identified as Latino/Hispanic. The participants' mean age was 36.13 years ($SD = 11.83$) and the range of ages was 25-58 years. All participants were teachers who have taught grades, ranging from preschool to twelfth grade. In terms of geographical location, seven participants were in Pennsylvania, two in Connecticut, two in Colorado and one each in Nevada, Washington, Oregon, and Maryland.

Measures

The same measures used in Pilot 1 were presented in Pilot 2. Four sets of information were collected: (a) demographic information, (b) TOK, (c) TOA, and (d) opinions on the video module and items presented.

Procedures

Similar to Pilot 1, the Pilot 2 participants were recruited through email and the social media website Facebook. The administration procedures of Pilot 2 were similar to Pilot 1. The main difference was the order of the measures. After reading an implied consent form, the teachers were randomly assigned to watch the ST video module or not. Then, teachers answered only the ST and classroom questions in the TOK (the multicultural teaching questions were removed). The order of the two sets of questions was randomly presented. Then all participants were randomly assigned to a scenario (ST or No-ST). After reading the scenario, participants answered the TOA. Then the participants answered the multicultural teaching set of questions, which is considered to be a section of the TOK. Finally, the participants completed the demographic questions. As part of the pilot process, participants were guided through the experiment again and asked to answer open- and closed-ended questions about the content and clarity of the video module, scenarios, and other items. Once participants had completed giving their feedback and to maintain their anonymity, participants were automatically redirected to a separate web link to register for the raffle. At the end of the data collection period for the second pilot, two \$10 electronic amazon gift cards were given to two randomly selected participants.

Results

Due to the small size of the sample of this pilot, the feedback from the participants was only examined qualitatively. In regard to the clarity of the instructions, items, and scenarios, none of the participants commented on lack of directions or expressed confusion on what was asked. Thus, adding instructions improved the clarity of the study. Two items in the classroom organization section of the TOK were changed due to the teachers' feedback. The item, "It is not important for teachers to evaluate if students have easy access to commonly used student tools

and resources,” was changed to “It is not important for teachers to ensure that students have easy access to commonly used student tools and resources (such as textbooks, worksheets, and other classroom materials).” The change was made, because a teacher’s response appeared to misinterpret the use of the word “evaluate.” Another teacher expressed that he or she was not sure what was meant by “student tools and resources,” so examples were included in the item to make the item clearer. The item, “When organizing a classroom, it is important for teachers to identify the best location to give instruction,” was changed to, “When organizing a classroom, it is important for teachers to identify the best location(s) to give instruction.” Three teachers found the item problematic, because there was usually more than one place ideal for instruction depending on the topic or subject and the type of instruction.

In terms of the ST questions in the TOK, only one item was changed to improve its clarity based on the teacher’s feedback. The item, “Students who experience chronic bouts of stereotype threat do not lose interest in the area relevant to the stereotype threat,” was changed to “Students who experience the threat of stereotype about a particular area over time may lose interest in that area.” A teacher responded that they were unsure about what “chronic bouts” meant so the wording was changed in an effort to make the item clearer. The teacher’s feedback also identified an unintended confound in the content of the scenarios. One teacher expressed confusion on whether the ST presented in the scenario was related to gender or socioeconomic status. As a result, the content of both scenarios were changed so that financial concerns were not noted to prevent socioeconomic status from becoming a confounding variable in the research design.

Finally, it was examined whether the video module was set up to effectively disseminate the ST information as intended. Multiple respondents ($n = 3$) reported that the video was too

long. Thus the video was shortened in two ways. The narration of the video was recorded with a slightly faster speaking rate. In addition, the summary portion of the video was eliminated. These changes reduced the length of the video from 6 minutes and 48 seconds to 5 minutes and 48 minutes. Finally, participants ($n = 2$) stated that there was too much text on the screen at one time, so the researcher eliminated non-essential text and shortened the text that appeared on screen when possible.

Other Revisions to the Research Design

Further examination of the placement of the measures resulted in additional changes to the sequence and presentation of materials. To mask the purpose of the study, it was decided not to present separately the multicultural teaching and ST questions. Instead, the TOK was modified. The ST and multicultural teaching items were reduced to five items each and were set up to be presented together, mixed among distractor items. The distractor items added were five items on reading instruction and five items on behavior management, so the final TOK would have 20 items related to the four following areas: ST, multicultural teaching, reading instruction, and behavior management. In addition, the classroom organization items were made into a separate measure. The distractor items are described in more detail in the measures section of the study. The TOK items are presented after the TOA in the final study. Another change made was in the demographic section about multicultural training. The concern continued to exist that asking about multicultural training would prime participants about the intent of the study. Instead of asking participants solely about multicultural training, another set of questions was included that focused on other work related training, such as reading instruction, math instruction, use of technology in the classroom, and response to intervention.

Actual Study

The actual study integrated the findings of Pilot 1 and Pilot 2 to test the stated hypotheses. The procedure and measures for the actual study are reviewed below.

Research Design

The research design of the study was a 2 x 2 experimental between subjects design, which involved two manipulations. The first manipulation was the intervention—exposure to an ST video module or not. The second manipulation was exposure to either a hypothetical ST scenario or not.

Knowledge module. Participants were randomly assigned to one of two module conditions. One condition exposed the participants to a video module on ST Knowledge. In the control condition, the pre-service teachers were not exposed to a video module. The ST Knowledge condition involved the presentation of a video, which lasted 5 minutes and 48 seconds. There were two segments in the video. The first segment of the video lasted 3 minutes and 10 seconds. In this segment, the video introduced the concept of ST and gave an overview of the situation in which ST is experienced (investment in performance, awareness of a stereotype, and acceptance of the stereotype), and then a short explanation of the importance of preventing ST is given. In the second segment, which lasted 2 minutes and 35 seconds, a review was provided of six strategies that teachers could use to prevent students from experiencing ST. The selection of these strategies was based on prior research (Good et al., 2008; Johns et al., 2008; Martens et al., 2006; Marx et al., 2009) on what has been found to be effective in preventing ST. For an in-depth review of the strategies and supporting research on ST, refer to the Prevention of ST section starting on page 17. A transcript of the content of the ST video module is presented in Appendix A.

Scenarios. The second manipulation was the type of hypothetical scenario presented to participants: ST scenario and No-ST scenario. The purpose of this manipulation was to test whether the exposure to ST knowledge was effective in teaching pre-service teachers about the concept and its application. The scenarios were brief and depicted a student who was having difficulties in school. In the ST scenario, the student was struggling in mathematics due to experiencing ST, evidenced by the student's (named Sophia) investment in the area of performance, the awareness of the relevant stereotype, and the awareness of confirming the stereotype (Steele & Aronson, 1995). Investment in the area of performance was operationalized by Sophia's comment about her desire to be a mathematician. Awareness of a gender stereotype is alluded to when Sophia noted that her brother says, "Girls can't do math." Her fear of the stereotype is confirmed in that (a) Sophia does not believe she has done well on the test and (b) everyone will think girls cannot do mathematics. The full scenario used in the study is contained in Appendix B.

In contrast, the No-ST scenario involved a student who lacked motivation in mathematics. Motivation was used as the subject matter, because Steele (1997) contended that lowered test performance differences between ST and control groups were due to the pressure of ST and not due to motivation. To ensure that this scenario did not contain any references to ST, several changes were made on investment in performance, awareness of stereotype, and fear of confirming the stereotype. In the No-ST scenario, Sophia is not invested in math, because she expresses that she does not need math. Sophia does not appear to be aware of ST, because information about related stereotypes are not mentioned and are purposely excluded. She also does not appear to be concerned or mention any concern about confirming any stereotypes. The No-ST scenario is presented in Appendix C.

Data Management of the Sample

A total of 102 pre-service teachers entered the portal to participate in the study. However, the number of actual participants was reduced to 85 due to missing or incomplete data. Data missing on 20% or more answers were from either the TOK ($n = 5$) or TOA ($n = 12$) measure; thus, these cases were removed. Most of the pre-service teachers from the target population were White (88.23%) and women (96.47%); this pattern was reflected in the sample. Three participants (3.52% of sample) identified as male; one participant identified as Black (1.18%), three as Asians (3.52%), four as Multiethnic (4.70%), and one as Latino/Hispanic (1.18%). This disproportion in the sample by gender and race/ethnicity could not be distributed across the experimental design. To maintain the integrity of the research design (no violations due to disproportional cell sizes based on demographics), these 12 cases, non-White and non-women, were removed from the sample. An additional 13 participants were removed from the sample, to ensure equal cell sizes and to minimize the violation of homogeneity of variance. Seven participants who had been exposed to the module and given the No-ST scenario and six participants who were not exposed to the module and were given the ST scenario were randomly removed, resulting in a final sample size of 60 White, women participants. The participants with missing data could not be compared to the sample, because the majority of their demographic data were also missing due to the order at which the items were presented. However, when the participants removed from the analyses to maintain balance are included in the sample the demographic data were found to be similar to the participants in the final sample.

Participants

The final sample size was 60 pre-service teachers attending a public, predominately, White university in the Mid-Atlantic region of the United States. Students ranged in age from 19

to 23 years ($M = 20.58$; $SD = 0.67$). The students' self-reported cumulative grade point averages ranged from 2.81 to 3.98 ($M = 3.61$; $SD = 0.25$). Their academic class level ranged from second year to fifth year in college. However, most participants were in their third year ($n = 53$; 88.3%), followed by fourth year ($n = 6$; 10.0%), and fifth year ($n = 1$; 1.7%). Over half of the participants reported their families' annual household income to be \$76,000 or higher ($n = 31$; 51.7%), followed by \$61,000 to \$75,000 ($n = 13$; 21.7%), then \$31,000 to \$45,000 ($n = 7$; 11.7%), then 46,000 to \$60,000 ($n = 6$; 10.0%), and \$16,000 to \$30,000 ($n = 3$; 5%). The students were also asked about their professional development opportunities in various areas that they have attended in the past five years. A summary of the frequency of participants' attendance to various professional development sessions is presented in Table 1.

Table 1

Frequencies of Participants' Professional Development on a Given Topic

Topic	Frequency of classes, workshops, or conferences attended			
	0	1-3	4-6	7-10
Math Instruction	33	25	0	2
Reading Instruction	24	31	3	2
Writing Instruction	28	29	1	2
Classroom Management	27	28	2	3
Response to Intervention	36	22	2	0
Special Education Related Topics	23	29	5	3
Diversity and Culture Related Topics	26	32	0	2
Use of Technology In the Classroom	39	19	0	1

Note. $N = 60$.

Measures

Four sets of measures were used to collect data: (a) demographic questions, (b) classroom organization questions, (c) interventions methods, and (d) a knowledge test about ST and teaching concepts (e.g., reading instruction, multicultural teaching, and behavior management).

Demographic information. Demographic information was collected on the following variables: race/ethnicity, gender, age, class standing, socioeconomic status, GPA, and number of conferences/classes/workshops attended on various topics related to teaching.

Classroom organization. The author created items about classroom organization, which focused on the best practices in teaching related to the organization of the classroom. The content of the items were based on material from McLeod et al. (2003), and Evertson (1989). An example of a classroom organization item is, “When organizing a classroom, it is important for teachers to identify the best location(s) to give instruction.” These questions are intended to be distractor items to shift the focus from ST. This measure contains five items. Respondents are asked to indicate whether each item is true or false. Thus, the correct options are rated as 1 and incorrect options as 0. Total score is based on summing across all the classroom organization items. Higher scores are considered to reflect a greater knowledge about classroom organization. The classroom organization questions are presented in Appendix D.

Test of Application. The TOA is used to assess participants’ ability to select appropriate strategies in response to a scenario. The TOA is composed of 11 possible school interventions teachers typically use with elementary students and is designed for use with the student in the hypothetical scenario. Participants are asked the following question: “As Sophia’s teacher what would you put in place to help Sophia achieve to what you think she is capable of in math?” The teachers then are asked to respond to the options for possible intervention (or lack of), rated on a scale from 1 (*Unlikely to Do*) to 5 (*Likely to Do*). Over half of the interventions are linked to ST (6 of 11), three interventions are from the ST video module, and three generalized from the ST condition. A higher rating is interpreted to mean that a teacher is more likely to use the prevention method. The final TOA is presented in Appendices B and C.

Test of Knowledge. The author also created the TOK for the study. The TOK is used to assess four knowledge areas (behavior management, reading instruction, multicultural teaching practices, and ST). The TOK contains 20 items, five questions each tapping four areas: (a) ST, (b) multicultural teaching, (c) reading instruction, and (d) behavior management. Respondents are asked to indicate whether each item is true or false. Thus, the correct options are rated as 1 and incorrect options as 0. Total score is based on summing across all items on the respective topic and dividing by five. Higher scores for the total scores are considered to reflect a beginning knowledge base in that area. The TOK is presented in Appendix E.

Procedures

The study was submitted and approved by the university's institutional review board as an exempt study. The author recruited pre-service teachers by seeking permission from instructors teaching curriculum courses. The author went to eleven literacy and math instruction courses for pre-service teachers and invited students to participate in the study. The courses that were targeted for recruitment and data collection were those required for pre-service teachers who were training to work with students from pre-kindergarten to fourth grade. In addition, only pre-service teachers with the major of Childhood and Early Adolescent Education were required to take the courses in which the recruitment of participants took place. On the same day of class recruitment, the instructor emailed a brief description and a link to the study, so those who wanted to participate could access the study through Qualtrics: Research Suite (2015). The description emailed to the participants is in Appendix F. Data collection took place during the spring and fall semesters of 2015. Students had a month to take part in the study and could complete it on a computer, tablet, or cell phone. No identifying information was collected in this study that could link the findings back to each student. Due to the nature of the design, a specific

sequence of the material was used after participants consented to participate in the study: (a) first, participants were randomly assigned the video module or not, (b) then, classroom organization questions were presented, (c) next, participants were randomly assigned to one of two scenarios, and (d) finally participants were administered the measures of TOA, TOK, and the demographic questions in the order presented. Within TOK, items representing the four knowledge areas (ST, multicultural teaching, behavior management, and reading instruction) were randomly mixed together. Completion of the study took approximately 10-20 minutes. After the study was completed, the pre-service teachers were redirected to another website in Qualtrics to redeem either 2 points extra credit from the course they were recruited from or a digital \$5 Amazon gift card. Of the 102 students who entered the portal to take part in the study, 79 students opted to receive extra credit ($n = 54$; 52.9%) or a gift card ($n = 25$; 24.5%).

RESULTS

Preliminary Analyses

Two sets of preliminary analyses were conducted prior to testing the hypotheses. The first set of analyses examined the factor structure of the measures created—TOA and TOK. The second set of analyses was conducted to check whether the demographic characteristics (i.e., age, GPA, class standing, and family's income) of participants were balanced or proportional across the conditions or not. SPSS (Version 22.0.0.0.) was used for all statistics.

EFA. EFAs were conducted on the scores for the TOA items. Scores for the items on the TOA met the assumptions for EFA (linearity, presence of small to moderate correlations [$|\lambda|$.01-.49]; $Mdn = |\lambda|.18$], and absence of multicollinearity ($r < .9$; Tabachnick & Fidell, 2001). One item, "Take no action" showed non-normality in the distribution of scores, exceeding the criteria of a skew of $|\lambda|3$ or less and a kurtosis of $|\lambda|8$ or less as acceptable (Kline, 2005). Means, standard deviations, skew, and kurtosis for TOA items are presented in Table 4. The correlation matrix for TOA items is also provided in the Appendix G.

The EFA was based on a principal axis extraction on the scores of the 11 TOA items, based on a sample size of 60. The correlation matrix was factorable (determinant of the matrix = .08; the Kaiser-Meyer-Olkin [KMO; Kaiser, 1974] = .58; Bartlett's test of sphericity [Bartlett, 1950]; $\chi^2(55) = 138.51, p = .001$). An oblique rotation (Direct Oblimin) was selected to allow for correlation between the factors (Hair et al., 2006).

Several criteria were used to determine the number of factors to retain: minimum average partial (MAP; Velicer, 1976); parallel analysis (PA; Horn, 1965), scree plot (Cattell, 1966); the size of the communality, simple structure (Thurstone, 1954), and the interpretability of the factor (Thompson & Daniel, 1996). The MAP procedure suggested that two factors should be retained,

but visual inspection of the scree plot (Cattell, 1966) indicated three factors. PA suggested five factors. Thus, several secondary EFAs were run based on the number of factors ranging from five to one.

The secondary solutions five through one were examined for which best reflected the 11 TOA items. Criteria used to select the best factor structure were (a) the number of items (minimum of 3) loading on each factor and (b) the magnitude of the pattern coefficients ($\geq |.4|$; Stevens, 2002). Four of the five solutions were not viable. The five-, four-, and three-factor structure had communalities with a wide range of values (e.g., .08 - .93), only two factors had three or more pattern coefficients $|.4|$ or higher, and simple structure was not supported. Furthermore, the five, four, and three factor solutions lacked theoretical convergence. For each of the factor solutions, at least one factor contained ineffective and effective prevention strategies. The one-factor solution was also rejected, because it accounted for less variance (17.32%) than the two-factor solution (29.53%). The two-factor solution appeared to be the most viable: (a) communalities ranged from .01 to .60 ($Mdn = .32$) and (b) each factor had three to five items with pattern coefficients ranging from $|.41|$ to $|.77|$ ($Mdn = |.59|$). However, two problems weakened the usefulness of this solution. One, Cronbach's alphas for the scores of both factors were less than .70: (a) .67 for Factor I and (b) .64 for Factor II. Two, the composition of items on each factor departed from the focus of the study—effective strategies for reducing ST. Items loading on the first factor reflected strategies that were more reactive in nature, meaning students would typically need to show some sort of difficulties to warrant actions by the teacher. On Factor II, salient items were negatively loaded and reflected strategies that could be considered to be more proactive in nature. Several items did not load on either factor or did not load saliently. Two items, “Peer model to work with” and “Emphasize

progress,” loaded on to the first factor but were not salient ($< .4$). One item, “Take no action,” did not load on either factor. An examination of the items revealed that items designated as effective ST strategies loaded on the same factor with those items identified as ineffective ST strategies. On the first factor, the salient ineffective methods of “Worksheets as practice” and “Tutoring during recess” loaded with the salient effective strategies of “Explain math tests as puzzles” and “Explain anxiety,” and the salient generalized effective strategy of “Students praise each other.” The second factor had three salient items, two ineffective ST prevention strategies (“Activities related to her interests” and “Reward her when she completes her work”), and one effective strategy (“Remind Sophia that she is a strong student”). As such, it was not possible to use these factors in the primary analyses to answer the research questions relevant to this study. Thus, the 11 TOK-II items were used individually in subsequent analyses. Factor loadings, communalities, eigenvalues, and reliability estimates of the two-factor solution are presented in Appendix H.

It was not possible to conduct a factor analysis on TOK or Classroom Organization because of the sample size and distribution of the scores. The TOK and Classroom Organization items are based on a nominal scale of measurement (true/false). There was an insufficient variability between the two binary responses to conduct a factor analysis. The majority of the participants (95% or more) answered 18 of the 20 items of the TOK as either true or false. Only two items, “When teaching young readers to comprehend text, it is not necessary to teach students to recognize text structure” (88.3%), and “When selecting teaching materials for reading, teachers do not need to select texts that specifically support their teaching goals” (78.3%), had less than 95% response rate of true or false. Because a factor analysis could not be conducted due to a lack of variability, the correlations were examined between the various scales of the TOK. A positive

relationship between the behavior management total and the reading instruction total was found to be statistically significant ($r = .41, p = .001$). Although related, these two variables were analyzed separately because they reflected distinct concepts.

For the classroom organization measure, there was also insufficient variability. Out of the five items, two items, “It will not be necessary for teachers to move freely and monitor students if the classroom is well organized” and “When organizing a classroom, it is important for teachers to identify the best location(s) to give instruction,” had limited variability in terms of the participants’ responses. A table of the participants’ response rates is presented in Appendix I.

Demographic variables and research design. Preliminary analyses were conducted to examine the distribution of pertinent demographic features: age, class standing, perceived household income, and GPA. Based on the two manipulations, the distributions of the demographic variables are shown in Tables 3 and 4. A priori alpha level was set at .01 and effect size was set at $\eta_p^2 = .16$, based on the effect size reported in a study about women’s ST experience related to learning (Rydell, Rydell, & Boucher, 2010). Two 2 x 2 ANOVAs were conducted to examine whether there were any statistically significant differences on age or GPA across the two manipulations. With age as the outcome variable, and module exposure and scenario exposure as the fixed factors, the 2 x 2 ANOVA was not statistically significant for any effect: (a) Module exposure [$F(1, 56) = 0.04, p = .85$], (b) Scenario exposure [$F(1, 56) = 0.32, p = .57$]; and (c) Module x Scenario effect [$F(1, 56) = 0.90, p = .35$].

The distribution of class standing, based on module and scenario, was examined and the finding is presented in Table 2. As noted before, most of the participants were in their third academic year.

Table 2
Percentage of Academic Class Standing based on the Conditions of Module and Scenario

Scenario	Module		Total	
	No ST video module	ST video module		
No-ST Scenario				
Academic Year				
Third	15	12	27	90.0%
Fourth	0	2	2	6.7%
Fifth	0	1	1	3.3%
Total	15	15	30	100.0%
ST Scenario				
Academic Year				
Third	11	15	26	86.7%
Fourth	4	0	4	13.3%
Fifth	0	0	0	0.0%
Total	15	15	30	100.0%
Sample Total				
Academic Year				
Third	26	27	53	88.3%
Fourth	4	2	6	10.0%
Fifth	0	1	1	1.7%
Total	30	30	60	100.0%

In regard to class standing, there were three students in their fourth and fifth years under the No-ST Scenario who received the ST video module, but there were no fourth and fifth years under the No-ST Scenario who received the control condition. This pattern was in reverse under the ST scenario. There were four students in their fourth year under the ST Scenario condition in the control condition, but no students beyond their third year in the ST video module condition who received the ST Scenario condition.

In regard to family income, there were variations in the participants' family income across the two manipulations, with more variations found under the ST scenario between the ST

video module and control conditions. Across all conditions, more participants reported family incomes \$76,000 or higher. A summary of this distribution of family income by manipulations, module and scenario, are presented in Table 3. Both class standing and family income have variability across the manipulations. However, class standing and family income cannot be used as covariates due to the disproportional cell sizes caused by the small sample size, so the findings should be interpreted with caution.

Table 3
Percentage of Family Income based on Conditions of Module and Scenario

Scenario	Module			
	No ST Video	ST Video	Total	
No ST Scenario				
Family Income				
\$16-30K	0	2	2	6.7%
\$31-45K	3	0	3	10.0%
\$46-60K	2	2	4	13.3%
\$61-75K	3	3	6	20.0%
\$76K or higher	7	8	15	50.0%
Total	15	15	30	100.0%
ST Scenario				
Family Income				
\$16-30K	0	1	1	3.3%
\$31-45K	2	2	4	13.3%
\$46-60K	1	1	2	6.7%
\$61-75K	2	5	7	23.3%
\$76K or higher	10	6	16	53.3%
Total	15	15	30	100.0%
Sample Total				
Family Income				
\$16-30K	0	3	3	5.0%
\$31-45K	5	2	7	11.6%
\$46-60K	3	3	6	10.0%
\$61-75K	5	8	13	21.7%
\$76K or higher	17	14	31	51.7%
Total	30	30	60	100.0%

Descriptive Statistics

The means, standard deviations, skew, and kurtosis for each item of the TOA and each of the TOK totals are reported in Table 4. The correlations between the variables are reported in Appendix G. The data were checked for meeting parametric assumptions. The assumption of

Table 4

Means, Standard Deviations, Skew and Kurtosis of the TOK and TOA

Variable	<i>M</i>	<i>SD</i>	Skew	Kurtosis
TOK				
1. Class Org Total	0.86	0.13	-0.38	-0.66
2. Behavior Total	0.97	0.08	-2.53	6.13
3. Read Total	0.92	0.12	-1.38	0.87
4. ST Total	0.61	0.06	0.77	7.55
5. MC Total	0.78	0.08	-2.52	10.00
TOA				
6. Take No Action	1.13	0.57	4.66	21.66
ST Prevention				
7. Explain Anxiety	3.87	1.14	-0.93	-0.07
8. Tests as Puzzles	3.35	1.05	-0.31	-0.36
9. Strong Student	4.68	0.62	-2.25	5.62
Generalized ST Prevention				
10. Student Praise	3.27	1.09	-0.15	-0.32
11. Peer Model	3.77	0.89	-0.71	0.67
12. Progress > Product	4.38	0.69	-0.68	-0.65
Ineffective ST Prevention				
13. Reward for Work	3.45	1.11	-0.45	-0.54
14. Task Interests	4.50	0.79	-1.80	3.08
15. Tutoring	2.67	1.22	0.15	-0.93
16. Additional Work	3.00	1.26	-0.05	-1.13

independent observations was met through the research design. The assumption of normality for the distributions of the scores was met (a skew of |3| or less and a kurtosis of |8| or less is acceptable; Kline, 2005), except for one item of the TOA and one of the total scores from the TOK. The average total score for Multicultural Teaching had a negative skew (-2.52), which fell

in the moderate range, but had a severe level of kurtosis (10). The average score for “Take no action” strategy on the TOA had a severe skew (4.66) and kurtosis (21.66). Given the small sample size of the sample, not obtaining a normal distribution of scores for all of the variables was expected. As such, results related to the Multicultural Teaching scale and the “Take no action” strategy will be interpreted with caution. Correlations between the TOA items ranged from $|.01|$ to $|.59|$ ($Mdn = |.12|$) and are presented in Appendix G. Reliability estimates for the scores of the TOA scales could not be calculated due to a lack of variability of scores. A table showing the frequency of responses for each item is presented in Appendix I.

Hypothesis 1

The first hypothesis was that pre-service teachers exposed to the ST video module would score higher on ST items than pre-service teachers who had not been exposed to the ST video module. Five 2 by 2 ANOVAs were conducted on each of the knowledge areas in TOK (behavior management, reading instruction, multicultural teaching, and ST) and classroom organization as the dependent variable, with module (ST exposure or no ST exposure) and scenario (ST or not) as the independent variables. For two of the ANOVAs, the assumption of homogeneity of variance was violated; the p-values were less than .01 for the Levene’s tests for the behavior management total ($p = .003$) and the reading instruction total ($p = .006$) from the TOK. However, ANOVAs are considered to be robust to minor or moderate violations of HOV when sample sizes are equal (Glass & Stanley, 1970), and the ratio between the smallest and largest variance is less than 1:3 (Dean & Voss, 1999). The ratios between the smallest and largest variance were less than 1:3, 21:100 and 23:100, respectively. All five 2 x 2 ANOVAs on classroom organization, behavior management, reading instruction, and ST knowledge based on

video module and scenario conditions were not statistically significant for main or interaction effects. A summary of the ANOVA findings is presented in Table 5.

Table 5

Summary of ANOVAs based on Module and Scenario on Classroom Organization and Four TOK Total Scores (N = 60)

Variables	SS	df	MS	F	p	η_p^2
Classroom Organization						
Scenario	0.024	1	0.024	1.47	.23	.03
Module	0.024	1	0.024	1.47	.23	.03
Scenario * Module	0.024	1	0.024	1.47	.23	.03
Error	0.912	56	0.016			
Total	0.984	60				
Behavior Management						
Scenario	0.000	1	0.000	0.00	1.00	.00
Module	0.000	1	0.000	0.00	1.00	.00
Scenario * Module	0.024	1	0.024	3.45	.07	.06
Error	0.389	56	0.007			
Total	0.413	60				
Reading Instruction						
Scenario	0.033	1	0.033	2.20	.14	.04
Module	0.006	1	0.006	0.40	.53	.01
Scenario * Module	0.017	1	0.017	1.12	.29	.02
Error	0.832	56	0.015			
Total	0.888	60				
Stereotype Threat						
Scenario	0.011	1	0.011	2.67	.11	.05
Module	0.003	1	0.003	0.67	.42	.01
Scenario * Module	0.000	1	0.000	0.00	1.00	.00
Error	0.224	56	0.004			
Total	0.238	60				
Multicultural Teaching						
Scenario	0.000	1	0.000	0.000	1.00	.00
Module	0.003	1	0.003	0.400	.53	.01
Scenario * Module	0.000	1	0.000	0.000	1.00	.00
Error	0.373	56	0.007			
Total	0.376	60				

Note. Effect size is η_p^2 = partial eta squared

Hypotheses 2

For Hypothesis 2, it was expected that pre-service teachers exposed to the ST video module would rate on average the ST strategies higher to use for the ST scenario in comparison to pre-service teachers who were not exposed to the ST video module. There were 11 strategies to be examined. To simplify the analyses and minimize Type I error, the strategies were grouped together in the following way: effective ST prevention strategies (“Explain Anxiety,” “Math Puzzles” and “Strong Student”), strategies generalized from ST prevention strategies (“Student Praise,” “Peer Model,” and “Emphasize Progress”), ineffective ST strategies (“Math Worksheets,” “Tutor at recess,” and “Interests”), and “No Action.” Three MANOVAs were conducted on each of the TOA groupings (effective strategies, generalized strategies, and ineffective strategies) as the dependent variable, with module (ST exposure or no ST exposure) and scenario (ST or not) as the independent variables, and class level and income as the covariates. In addition, a one-way ANOVA was conducted, with “No Action,” as the dependent variable and the module and scenarios as the factors.

For the MANOVA on ST strategies was conducted, the Box’s M was statistically significant, Box’s M = 62.63, $F(18,11081.83) = 3.14$, $p = .001$), violating the assumption of homogeneity of variance-covariance matrices. However, Box’s M is highly sensitive and it is suggested that MANOVAs are robust when the Box’s M is found to be statistically significant and the sample sizes are equal (Tabachnick & Fidell, 2001). The MANOVA was only practically and statistically significant for the main effect of Scenario (ST or not) on ST strategies, Wilk’s $\Lambda = .57$, $F(3, 54) = 13.09$, $p < .001$, $\eta_p^2 = .43$. A summary of the MANOVA is provided in Table 6.

Table 6

Summary of MANOVA of Module Exposure and Scenario Conditions on ST Strategies (Explain Anxiety, Math Puzzles, and Remind Student; N = 60)

Variables	Wilk's Λ	F	Hypothesis Df	Error df	p	η_p^2
Scenario	0.57	13.81	3	53	.01	.43
Module	0.93	1.31	3	53	.28	.07
Module*Scenario	0.97	0.63	3	53	.60	.03

A descriptive discriminant analysis (DDA) was conducted to interpret the linear composite of the three ST prevention strategies. The overall model was statistically significant, with the first discriminant function accounting for 100% of the variance, Wilk's $\Lambda = .573$, $\chi^2(3) = 31.43$, $p < .001$. The standardized function coefficients ranged from -0.01 to 1.05, with Explain Anxiety (1.05) contributing the most to the function, followed by Math Puzzles (-0.18) and Remind Student (-0.01). The canonical variate structure coefficients ranged from .09 to .99, with Explain Anxiety accounting for approximately 97% of variance in the variate, followed by Math Puzzles ($s = .21$) and Remind Student ($s = .09$). The variate was named *Explain Anxiety*. On average, the ST strategy Explain Anxiety best discriminated the pre-service teachers in the ST scenario group (0.85) from the pre-service teachers in the no-ST scenario group (-0.85). The centroids of this function are presented in Figure 1, and a summary of the structure and discriminant function coefficients is provided in Table 7.

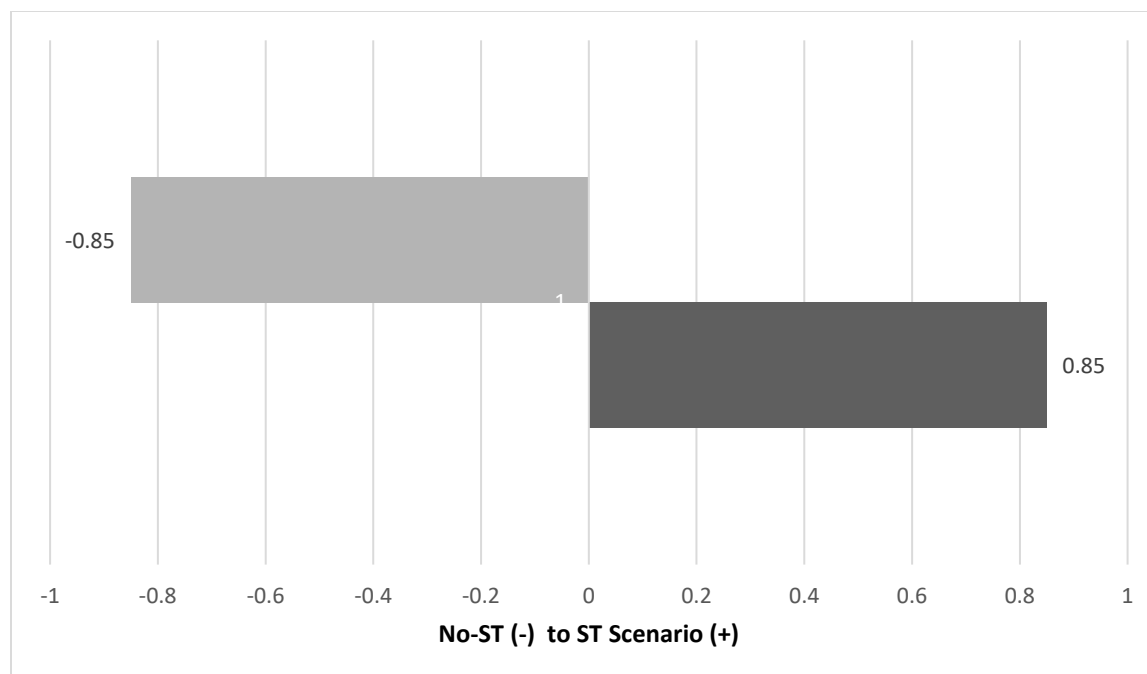


Figure 1. Graphical depiction of the centroids for the discriminant function of the ST prevention strategies ($N = 60$). The variate was labeled “Explaining Anxiety,” showing those in the ST Scenario condition (positive direction) were more likely than those in the No-ST Scenario condition (negative direction) to use this prevention strategy.

Table 7

Standardized Function and Structure Coefficients for Descriptive Discriminant Analyses for the ST Prevention Strategies Based on the Scenario Conditions ($N = 60$)

	Standardized Function Coefficients	Structure Coefficients
Explain Anxiety	1.05	0.99
Math Puzzles	-0.18	0.21
Remind Student	-0.01	0.09

For the MANOVA on the ST generalized prevention strategies, Box’s M was not statistically significant, Box’s $M = 43.50$, $F(18,11081.82) = 1.505$, $p = .08$), supporting the assumption of homogeneity of variance-covariance matrices. The omnibus MANOVA was not statistically significant for the interaction effect or main effects. Thus, no further analyses were conducted. A summary of the MANOVA is provided in Table 11.

Table 8

Summary of MANOVA of Module and Scenario Conditions on Generalized Prevention Strategies (N = 60)

Variables	Wilk's Λ	F	Hypothesis Df	Error df	p	η_p^2
Scenario	0.90	1.92	3	54	.14	0.10
Module	0.94	1.15	3	54	.34	0.06
Module * Scenario	0.90	1.88	3	54	.144	0.10

Note. Effect size is η_p^2 = partial eta squared

For the MANOVA on the set of ineffective strategies, the Box's M was not statistically significant, Box's M = 31.99, $F(30,8622.13) = 0.93$, $p = .58$), supporting the assumption of homogeneity of variance-covariance matrices. The MANOVA was not statistically significant for the interaction or the main effects. Thus, no further analyses were conducted. A summary of the MANOVA is provided in Table 9.

Table 9

Summary of MANOVA based on Module and Scenario for the Set of Ineffective Strategies (N= 60)

Variables	Wilk's Λ	F	Hypothesis df	Error df	p	η_p^2
Scenario	0.80	3.36	4	53	.02	.20
Module	0.98	0.22	4	53	.93	.02
Interaction of Scenario and Module * Scenario	0.99	0.15	4	53	.96	.01

Note. Effect size is η_p^2 = partial eta squared. The alpha level was set a priori at $p < .01$.

In regard to the ANOVA with "No Action" as the dependent variable, and scenario and module as the independent variables, it was not statistically significant. A summary of these results is presented in Table 10.

Table 10

Summary of One-Way ANOVA based on Module and Scenario Conditions on the TOA Item Take No Action (N = 60)

Variables	SS	df	MS	F	p	η_p^2
Scenario	0.067	1	0.067	0.21	.65	.00
Module	1.067	1	1.067	3.37	.07	.06
Scenario * Module	0.067	1	0.067	0.21	.65	.00
Error	17.733	56	0.317			
Total	18.934	60				

Note. Effect size is η_p^2 = partial eta squared. The alpha level was set a priori at $p < .01$.

DISCUSSION

The purpose of this study was to examine whether exposing pre-service teachers to a brief training video on ST would increase their knowledge about ST and to apply such information to a hypothetical student. Two hypotheses were posed. The first focused on whether pre-service teacher's knowledge of ST would be higher when exposed to a short lesson on ST in comparison to those who were not, and the second focused on whether pre-service teachers exposed to ST video were likely to rate ST prevention strategies as more appropriate to work with student described experiencing ST in comparison to those who were not exposed to the video but to the ST case. Neither of the hypotheses was supported. An examination of the findings, or the lack thereof, is addressed by hypothesis. Then limitations, future directions, and implications are addressed.

Hypothesis 1: Exposure to ST Video Module and Knowledge

It was hypothesized that pre-service teachers who watched the ST video module would score higher on the ST knowledge test than pre-service teachers who had not watched the ST video module. The hypothesis was not supported: Pre-service teachers' ST knowledge scores did not differ between those who watched the ST video module and the control group. There are several possible explanations for the lack of significance for the first hypothesis. The most obvious explanation is that pre-service teachers are more aware of concepts related to ST than expected. The vast majority of the participants ($n = 58$), regardless of their experimental condition answered 3 out of the 5 ST knowledge questions correctly. This finding indicates that both groups may have already had some familiarity with ST or concepts related to ST. In addition, over half of the participants (34 participants) indicated that they had attended one or more classes, workshops, or conferences dedicated to diversity and culture related topics. This

information may indicate an increase in the availability of multicultural courses and interest in multicultural issues. Previous research has showed that teachers reported little to no training in multicultural education (Sharma, 2005). Thus, pre-service teachers may be more knowledgeable, because of their exposure to and interests in multicultural concepts through course work and professional training.

Another explanation for the findings is the methods of measurement used. It is possible that the method of measuring the participant's knowledge of ST was not appropriate. In the current study, the items measuring ST knowledge were presented in a true or false (TF) format, which may not have detected the nuanced differences in knowledge between the control group and the group exposed to the ST video module. Dichotomous response options (e.g., true-false, yes-no) are not ideal to measure knowledge, and more response options, i.e., multiple choice items, are more likely to result in scores that better represent participant knowledge (Agble, 1999). This TF format of questions is more susceptible to guessing the correct answer than a multiple-choice format (Brame, 2013). In addition, research found that response categories of 7, 8, 9, or 10 response options resulted in the most reliable scores (Preston & Coleman, 2000). In terms of the current study, the TF questions also resulted in a lack of variation. Many of the participants responded similarly to the items. For all of the ST items, 95% or more of the participants responded with the same response. For example, ST Items 2 and 3 both had 100% of participants answered false and Item 4 had 95% of respondents select false. ST Items 1 and 5 had 98.3% and 96.7% of participants responding true, respectively. A larger sample size may have increased the variation in responses, if the issue was the format, not the knowledge level of the participants. An additional problem beyond the TF format was the difficulty of the questions. Most of the participants (96.67%), regardless of experimental condition, obtained a score of three

out of five correct, an indication that the questions were too easy for these pre-service teachers. Either these participants were all fairly knowledgeable about ST or the questions reflected the easiest or obvious aspects of the ST construct, which could be discerned by a reasonably intelligent person, i.e., bachelor's level education students. For example, "Only females experience the threat of stereotypes" had a 100% correct response rate from the participants. The TOK items did not adequately differentiate between those who were exposed to the ST video module and those who were in the control condition. It appears that questions with a greater level of difficulty (e.g., 75% of participants responded correctly; Thompson & Levitov, 1985) may have allowed for more variation in responses and may have allowed for differences between groups (e.g., knowledge differences between those who were exposed to the ST video module and those who were not). Although a series of pilot studies were conducted to ensure that the design of the study and the questions were satisfactory, the focus was not on item difficulty.

Hypotheses 2: ST Video Module Exposure and Selected Prevention Strategies

For Hypothesis 2, it was expected that pre-service teachers exposed to the ST video module on average would rate a higher preference to using ST strategies for the ST scenario than pre-service teachers in the control condition would rate ST strategies for the ST scenario. This hypothesis was not supported. There are several explanations for why the second hypothesis was not supported.

One possible reason for the nonsignificant findings may have been due to the manipulation itself—the ST video module. A large portion of the ST video module described prevention strategies for ST (2 minutes and 35 seconds). However, the ST video module did not enable teachers to participate in active learning. Research has supported that part of an effective

professional development is active learning, such as engaging in discussion with other professionals (Birman, Desimone, Porter & Garet, 2000).

It is also possible that pre-service teachers picked socially desirable responses rather than indicating their actual likelihood to use various ST prevention strategies. Socially desirable responses refers to participants making themselves appear more positively through their answering of questions in terms of prevailing cultural norms (Krumpal, 2013). To determine whether socially desirable responses had an impact on the results, questions testing whether participants were responding in socially desirable ways should have been included (Krumpal, 2013). However, social desirability questions were not embedded in the study, and therefore, the influence of social desirability on the findings is unknown.

Although the second hypothesis was not supported, there was an unexpected finding. A main effect for the scenario manipulation was found. Pre-service teachers exposed to the ST scenario on average preferred to use the ST prevention strategy of explaining the anxiety to the student in comparison to those pre-service teachers exposed to the No-ST scenario. This effect may have been due to a confounding variable in the ST scenario. Participants exposed to both the ST video module and scenario may have been reacting to the ST scenario by selecting an appropriate ST prevention strategy learned from the ST video module. However, participants not exposed to the ST video module reacted in a similar manner to the ST scenario. It is possible that the participants not exposed to the ST video module interpreted the ST scenario reflective of another theoretical base. The similarities between ST and other theories has been a critique of ST and some believe that ST may be better explained by other theories, such as test anxiety, specifically Yerkes-Dodson law (Jensen, 1998). Jensen (1998) believed that the effect of ST could also be explained through the interaction effect between ability and test complexity on test

anxiety. However, others have questioned that the Yerkes-Dodson law better explains ST, because task complexity in the Yerkes-Dodson law has not been well defined (Diamond et al., 2007). In the current study, the participants not exposed to the ST video module may have believed that the student in the ST scenario was experiencing test anxiety and they were attempting to intervene appropriately to assist a student with test anxiety. Test anxiety is defined as negative feelings of tension and worry when an individual faces a demanding evaluative situation (Spielberger, 1972). ST can also occur in evaluative situations, when the situation is also considered to be challenging (Steele & Aronson, 1995). The ST scenario included information that may have led participants who were not exposed to the ST video module to believe that the student may have been experiencing test anxiety. For example, the ST scenario mentioned that the student's scores on tests were lower than on assignments and that the student expressed worry in relation to wanting to do well on her tests. In addition, explaining anxiety, which was found to be effective for students who were experiencing ST (Johns, Inzlicht, & Schmader, 2008; Johns, Schmader, & Martens, 2005), has also been found as an effective intervention for students experiencing test anxiety. Sarason (1981) found that encouraging student discussion about their concerns and providing support to students with high feelings of test anxiety resulted in higher scores on a challenging task for these students in comparison to similar students who did not receive social support.

Limitations

The issues addressed above provide possible explanations for the findings as well as highlighting the limitations to the study. Although several pilots were conducted to ensure the integrity of the ST video module, the pilots seemed to have been insufficient in addressing the potential problems identified in the current study—the amount of time of video and its focus. For

example, the ST video module gave an example of ST and the different components of the construct, but it did not provide a review of No-ST examples. The participants who viewed the ST video module could have benefited from examples of No-ST conditions, given that the task of recognizing examples of ST and No-ST conditions was inherent in the Test of Application. Additionally, there were other limitations that threatened the internal and external validity of the study. The results of this study may not be generalizable, because the pre-service teachers were only recruited from one university in a primarily rural area in a Mid-Atlantic state in the US. In addition, participants were difficult to recruit for the final sample. There were a limited number of students each semester who were eligible to take part in the study, in that new participants were taking classes with students who had already participated in the study. Timing was also a critical factor, because recruitment had to be negotiated around course exams, field outings, and guest speakers. Some course instructors were unable to accommodate the recruitment of students during class time. This constraint limited opportunities for recruitment, and possibly limited the sample size and range of data. Another limitation was the number of participants in each cell when conducting analyses. Unequal and small cell sizes make using parametric statistics, such as ANOVA and MANOVA, challenging, as violation to homogeneity of variance and covariance matrix becomes a given. In addition, unequal or small cell size can result in inflated risk of Type I error and a decrease of statistical power (Warner, 2012).

Another limitation was the use of one item to examine previous exposure to multicultural teaching coursework. The item asked pre-service teachers to indicate how many classes, workshops, or conferences attended. However, a weekend long workshop and a semester long course are different. It may have been better to use several questions to tap into previous multicultural training so previous experience could have been better examined.

Future Directions

One future direction that emerged due to the limitations of the current study is to use a larger sample size. The minimum sample size found to detect a statistical significant as well as meaningful finding was $n = 128$, from a power analysis using G*Power version 3.1.0, based on the following criteria: (a) the effect size set at .25 (η^2), based on the Rydell et al.'s (2009) ST intervention findings, (b) power level set at .80; and (c) alpha level set at .05. The larger sample size would result in larger cell sizes of the factors for using the statistics of ANOVA and MANOVA, which would likely reduce the risk of a Type I error. A larger and more heterogeneous sample may also provide increased variability in characteristics, such as age and training across the factors. In addition, it is recommended that pre-service teachers be recruited from more than one university to increase the generalizability of the findings.

Two, another direction for future research is to examine ST and the application of interventions in naturalized settings. Much of the ST intervention research has been conducted in a laboratory setting at a university (e.g., Ambady et al., 2004; Martens et al., 2006; Quinn & Spencer, 2001). As a result, it has been unclear whether the findings in such a setting can be generalized to the field. Future directions of study could include a case study to examine how ST is invoked and intervened or a long-term structured observation could be used to examine how interactions within the classroom invoke as well as prevent students experiencing ST.

Three, the ST video module and scenarios designed to educate pre-service teachers need to be further researched as viable methods for training teachers. Research suggests that video length is one of the most important factors in engaging the viewers, recommending that online videos be kept to less than 6 minutes in length (Guo, Kim & Rubin, 2014). Although the ST video module used in this research was less than 6 minutes, other factors could have affected pre-

service teachers' learning of the material. For instance, research has shown that a video of a person talking is more engaging than other media like voice over with text (Guo et al., 2014). Thus, a future direction may be to create another module on ST that features a person or different persons presenting the material to examine its effectiveness.

Four, it may be beneficial to examine pre-service teachers' preferences for ST prevention strategies. In the current study, pre-service teachers were likely to use the "explaining anxiety" prevention strategy in the ST scenario, regardless of exposure to the ST video module. Research dedicated to why teachers select methods of ST prevention may help researchers be better able to target prevention strategies to examine that are more likely to be used by pre-service teachers.

Five, due to the unexpected finding and possible confound of test anxiety, it may be beneficial to conduct a similar study to identify and eliminate possible confounding variables. Additional research with measures dedicated to examining participant knowledge related to test anxiety and appropriate interventions could aid in further refinement of the design and intervention during piloting phases. In addition, it may be helpful to examine teachers' responses to a test anxiety scenario, in addition to the No-ST and the ST scenario.

Six, the current study attempts to be the first step in having teachers to address a stigma that students may encounter. However, many intrapersonal considerations may greatly change the effectiveness of the professional developmental activity. ST Research may be useful in comparing pre-service teachers and practicing teachers. Practicing teachers may respond differently, i.e., selecting prevention strategies and reacting differently to the same professional development opportunity, than pre-service teachers. Differences in responses could be due to years in teachers and the types of student-teacher interactions encountered. In addition, another direction for research would be to use standardized measures (e.g., Teacher Multicultural

Attitude Survey; Ponterotto, Baluch, Greig, & Rivera, 1998; & Teachers' Perception of Risk Scale; Teo & Fevre, 2016) to examine teacher attitudes and beliefs on multicultural awareness and perception of risk in making educational changes through pre- and post-assessment.

Seven, another area of research of future research would be to examine the requirements for an effective ST interventionist. Future research could examine the effectiveness of the intervention if an interventionist is the person invoking the ST in the participant. Other interventionist traits could also be examined, such as does a power differential between the student and interventionist affect effectiveness and what is the ideal interventionist to student ratio for intervening on ST?

Eight, another future direction for research would be to continue to conduct Wise Interventions related to ST prevention. Further research on interventions that address the underlying psychological processes that cause ST may be useful. In addition, examining other methods of long lasting interventions of ST may be useful because only the incremental view of intelligence approach to ST prevention has empirical support for its long term benefits (Cook et al., 2012).

Implications

To date, no other study has been conducted to examine pre-service teachers' response to a module on ST and its prevention. This study provides a framework for how similar studies could be conducted. Findings from similar studies could help identify components that will make the module on ST more effective in training pre-service teachers, and further research could help refine the intervention by identifying what components are necessary to ensure optimal learning. Although the ST video module in the current study did not support the hypotheses, the lack of support may be due to methodological shortcomings. Finding effective ways to teach pre-service

teachers about ST and prevention methods are important considering that many teachers feel unprepared to teach a diverse classroom (National Center for Education Statistics, 2000). Previous research has showed that the integration of multicultural training into teacher training programs is lacking (Gollnick, 1995). However, this lack of multicultural training may no longer be the case given the absence of difference in ST knowledge between the pre-service teachers exposed to the module versus those not exposed. In essence, this study may indicate that pre-service teachers may be gaining some basic knowledge of ST from their multicultural related course content.

Literature supports the existence of ST and that it has been activated through various stereotypes, such as race (Steele & Aronson, 1995), gender (Nguyen & Ryan, 2008), and age (Hess, Auman, Colcombe, & Rahhal, 2003), resulting in decreased performances in various domains like memory (Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005), academics (Steele & Aronson, 1995), driving (Yeung & von Hippel, 2008), sports (Hermann & Vollmeyer, 2016), and work place (Bergeron, Block, & Echtenkamp, 2006). Research has also been conducted to examine various methods of ST prevention (Marx et al., 2009; Quinn & Spencer, 2001; Rydell et al., 2009; Schimel et al., 2004), and this study contains the most comprehensive review of empirically supported ST prevention strategies. The prevention strategies compiled in this study could be helpful to researchers who are interested in examining ST prevention strategies. The literature review of this study may be used as a foundation for researchers to move from studying the existence of ST to more clinical applications (e.g., standardized prevention procedures for teachers and training programs) that could aid with ST prevention.

In terms of the implications for school psychologists, there are factors to consider in regard to working with students who may be experiencing ST. During psychoeducational

testing, Jordan and Lovett (2006) suggest that school psychologist should try to avoid topics of conversation that may unintentionally activate a group membership for initial rapport building. For instance, avoid talking about shopping (may activate socioeconomic status) and music preferences (may activate socioeconomic status, ethnic groups, or regional affiliations). In addition, based on previous ST research (Ambady et al., 2004; Steele & Aronson, 1995), school psychologists may also want to interview students after ability and achievement testing are completed to minimize activating any identities, because typically interviewing will tap into things like home environment, family background, and expectations placed upon the student as well as self-perceptions of competency. School psychologists will also want to avoid describing psychoeducational tests like achievement and intellectual ability tests as diagnostic in nature, because as Steele and Aronson (1995) found task description can greatly affect whether ST is experienced. Given concerns related to ST, it is also important that school psychologists consider ST when interpreting test results. School psychologists may want to ask informal interview questions related to ST or give measures of state or test anxiety to try to assess whether ST was experienced during the testing.

Going into classrooms and speaking to students about stereotype threat and normalizing that it is okay to feel nervous, especially during important tests, may benefit students. This approach may prevent them from experiencing stereotype threat, because providing explanations of ST and the accompanying feeling can prevent students from experiencing ST (Johns, Inzlicht, & Schmader, 2008; Johns, Schmader, & Martens, 2005). School psychologists can also help prevent ST by using language that fosters a belief in incremental intelligence and encouraging teachers and staff within the schools to begin using such language as well, because a view of

incremental intelligence has been found to prevent the effects of ST (Aronson et al., 2002; Good et al., 2003).

School psychologists can also assist teachers by providing in-service and professional developmental opportunities related to multicultural education. Multicultural education will benefit the students and work to prevent ST in several ways. For instance, the use of role models has been seen in research to protect students from experiences ST (Marx et al., 2009; McIntyre et al., 2003), so encouraging teachers to include role models who may be counter stereotypical into lessons and class discussions may be helpful. In addition, to minimize ST, students may be better able to identify that they have multiple identities and that they are individuals rather than a representative of a group to which they belong (Ambady et al., 2004).

Conclusion

ST is an important multicultural phenomenon that has implications for teachers and the increasingly diverse student population in US classrooms (Campbell, 2009). The purpose of the study was to examine pre-service teachers' response to a ST video module as well as their self-reported likelihood of implementing prevention strategies when given a scenario about a student experiencing ST. None of the hypotheses were supported, possibly due to methodological shortcomings. However, this study provides a foundational step toward researching teacher's ability to select and implement ST prevention strategies within a classroom setting. This study may serve as a template for examining teacher's knowledge of and use of ST prevention strategies. Future research is recommended to continue to examine ways to effectively disseminate information to pre-service teachers and teachers about ST and prevention strategies.

REFERENCES

- Agble, P. K. (1999). *A psychometric analysis of different scoring strategies in statistics assessment*. Unpublished doctoral dissertation, Kent State University.
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley Publishing Company.
- 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, 401-408. doi: 10.1016/j.jesp.2003.08.003
- Alter, A., Aronson, J., Darley, J., Rodriguez, C., & Ruble, D. N. (2009). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology*, 46, 166-171. doi: 10.1016/j.jesp.2009.09.014
- Aronson, J. (2009). *Improving academic achievement: Impact of psychological factors on education*. New York, NY: Academic 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, 113-125. doi: 10.1006/jesp.2001.1491
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology*, 4, 1-2.
- Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviations, and spillover. *Journal of Experimental Psychology* 136, 256-276. doi: 10.1037/0096-3445.136.2.256

- Bergeron, D. M., Block, C. J. & Echtenkamp, B. A. (2006). Disabling the able: Stereotype threat and women's work performance. *Human Performance*, 19, 133-158. doi: 10.1207/s15327043hup1902_3
- Birman, B., Desimone, Garet, M., & Porter, A. (2000). Designing professional development that works. *Educational Leadership*, 57(8), 28-33.
- Blair, A. N., & Steele, J. R. (2010). Examining the consequences of exposure to racism for the executive functioning of Black students. *Journal of Experimental Social Psychology*, 46, 127-132. doi: 10.1016/j.jesp.2009.08.016
- Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2013). Enhancing a brief writing intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology*, 105(2), 427.
- Brame, C., (2013). *Writing good multiple choice test questions*. Retrieved [7/10/2017] from <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>.
- Buunk, B. P., & Ybema, J. F. (1997). Social comparisons and occupational stress: The identification-contrast model. In B. P. Buunk, & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 359-388). Mahwah, NJ: Erlbaum.
- Campbell, D. E. (2009). *Choosing democracy: A practical guide to multicultural education*, (4th ed.). Upper Saddle, NJ: Prentice-Hall.
- Carter, N. (1991). Learning to measure performance: The use of indicators in organizations. *Public Administration*, 69, 85-101. doi: 10.1111/j.1467-9299.1991.tb00783.x
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276. doi: 10.1207/s15327906mbr0102_10

- Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Aging Research*, 31, 235-260. doi: 10.1080/03610730590948177
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313, 1307-1310. doi: 10.1126/science.1128317
https://ed.stanford.edu/sites/default/files/recursive_processes_final_science.pdf
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Raising minority performance with a values-affirmation intervention: A two-year follow-up. *Science*, 324, 400-403. Retrieved 7/17/17 from
https://ed.stanford.edu/sites/default/files/recursive_processes_final_science.pdf
- Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: Protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology*, 102, 479–496. doi: 10.1037/a0026312
- Croizet, J., & 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, 588-594.
doi: 10.1177/0146167298246003
- Croizet, J. C., Dutrévis, M., & Désert, M. (2002). Why do students holding non prestigious high school degrees underachieve at the university? *Swiss Journal of Psychology*, 61, 167-175. doi: 10.1024//1421-0185.61.3.167

- Darley, J. M. & Latane, B. (1986). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377-383.
doi:10.1037/h0025589
- Davis, N., Preston, C. & Sahin, I. (2009). Training teachers to use new technologies impacts multiple ecologies: Evidence from a national initiative. *British Journal of Educational Technology*, 40, 861–878. doi:10.1111/j.1467-8535.2008.00875.x
- Dean, A. and Voss, D. (1999). *Design and analysis of experiments*. New York: Springer.
- Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J. & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson Law. *Neural Plasticity*, ID 60803. doi:10.1155/2007/60803
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048. doi: 10.1037/0003-066X.41.10.1040
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Evertson, C. M. (1989). Improving classroom management: A school-based program for beginning the year. *Journal of Educational Research*, 83, 173-188. doi: 10.1080/00220671.1989.10885935
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Forbes, C. E., Schmader, T., & Allen, J. J. B. (2008). The role of devaluing and discounting in performance monitoring: A neurophysiological study of minorities under threat. *Social Cognitive and Affective Neuroscience*, 3, 253-261. doi: 10.1093/scan/nsn012

- Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ, Prentice-Hall.
- Gollnick, D. (1995). National and state initiatives for multicultural education. In J. Banks, & C. Banks (Eds.), *Handbook of research on multicultural education*. Old Tappan, NJ: Macmillan.
- 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, 659-670. doi: 10.1177/0146167202288010
- Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, 29, 17-28. doi: 10.1016/j.appdev.2007.10.004.
- 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, 645-662. doi: 10.1016/j.appdev.2003.09.002
- Griffin, B. W. (2002). Academic disidentification, race, and high school dropouts. *High School Journal*, 85, 71-81. doi: 10.1353/hsj.2002.0008
- Guo, P. J., Kim, J., & Rubin, R. (2014, March). *How video production affects student engagement: An empirical study of MOOC videos*. Paper presented at the first ACM Learning@ scale conference. Retrieved 7/17/17 from <http://up.csail.mit.edu/other-pubs/las2014-pguo-engagement.pdf>. doi: 10.1145/2556325.2566239

- Guskey, T. R. (2002a). Professional development and teacher change. *Teachers and Teaching: Theory and Practice*, 8, 381-391. doi: 10.1080/13540600210000051_2
- Guskey, T. R. (2002b). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), 45-51. Retrieved from <http://tguskey.com/wp-content/uploads/Professional-Learning-4-Evaluating-Professional-Development.pdf>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, Rolph E., & Tatham, R. L. (2006). *Multivariate Data Analysis* (6th ed.), NJ: Pearson Education, Inc.
- Hermann, J. M., & Vollmeyer, R. (2016). “Girls should cook. Rather than kick!”- Female soccer players under stereotype threat. *Psychology of Sport and Exercise*, 26, 94-101. doi: 10.1016/j.psychsport.2016.06.0101469-0292/
- Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *Journal of Gerontology: Psychological Sciences*, 58, 3-11. doi: 10.1093/geronb/58.1.P3
- Hess, T. M., Hinson, J., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults’ memory performance. *Experimental Aging Research*, 35, 153-177. doi: 10.1080/03610730802716413
- Hodges, M., Hodges Kulinna, P., Lee, C., & Kwon, Ja Youn. (2017). Professional development and teacher perceptions of experiences teaching health-related fitness knowledge. *Journal Teaching in Physical Education*. 36, 32-39. doi: 10.1123/jtpe.2016-0107
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30: 179–185.

- 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, 365-371. doi: 10.1111/1467-9280.00272
- Inzlicht, M., Tullett, A. M., Legault, L., & Kang, S. K. (2011). Lingering effects: Stereotype threat hurts more than you think. *Social Issues and Policy Review*, 5, 227-256. doi: 10.1111/j.1751-2409.2011.01031.x
- James, W. (1890). *The principles of psychology*. NY: Holt.
- Jensen, A. R. (1998). Population differences in *g*: causal hypotheses in Jensen, A. R. (Ed.), *The g factor: the science of mental ability*, Westport, CT: Praeger.
- Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology*, 137, 691-705. doi: 10.1037/a0013834
- 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, 175-179. doi: 0.1111/j.0956-7976.2005.00799.x
- Jordan, A. H., & Lovett, B.J. (1990). Stereotype threat and test performance: A primer for school psychologists. *The Journal of School Psychology*, 45, 45-59. doi: 10.1016/j.jsp.2006.09.003
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36. doi: 10.1007/BF02291575
- Katz, I., Roberts, S., & 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. doi: 10.1037/h0022080

- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology*, 77, 323-338. doi: 10.1348/000709906X113662
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47, 2025-2047. doi: 10.1007/s11135-011-9640-9
- Lazarus, R.S. (1991). *Emotion and adaptation*. Oxford, UK: Oxford University Press.
- Levy, B. (1996). Improving memory in old age through implicit self stereotyping. *Journal of Personality and Social Psychology*, 71, 1092-1107. doi: 10.1037/0022-3514.71.6.1092
- Lockwood, P. (2006). Someone like me can be successful: Do college students need same-gender role models? *Psychology of Women Quarterly*, 30, 36-46. doi: 10.1111/j.1471-6402.2006.00260.x
- Martens, A., Johns, M., Greenberg, J., & Schimel (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, 42, 236-243. doi: 10.1016/j.jesp.2005.04.010
- Martines, D. (2008). *Multicultural school psychology competencies: A practical guide*. Thousand Oaks, CA: Sage Publications.
- Marx, D. M., Ko, S. J., & Friedman, R. A. (2009). The “Obama effect”: How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology*. 45, 953-956. doi: 10.1016/j.jesp.2009.03.012.

- Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology, 44*, 645-657. doi: 10.1348/014466604X17948
- Marx, D. M. & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*, 1183-1193. doi: 10.1177/01461672022812004
- McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486-493. doi: 10.1016/j.appdev.2006.06.003.
- McIntyre, R. B., Paulson, R., & Lord, C. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*, 83-90. doi: 10.1016/S0022-1031(02)00513-9
- McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Development, 74*, 498–515.
- McLeod, J., Fisher, J., & Hoover, G. (2003). *The key elements of classroom management: Managing time and space, student behavior, and instructional strategies*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Muraven, M. R., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*, 247-259. doi: 10.1037//0033-2909.126.2.247
- Najdowski, C. J., Bottoms, B. L., & Goff, P. A. (2015). Stereotype threat and racial differences in citizens' experiences of police encounters. *Law and Human Behavior, 39*, 463–477. doi: 10.1037/lhb0000140

- National Center for Education Statistics. (2000). *Teacher preparation and professional development: 2000*. Retrieved from <http://nces.ed.gov/surveys/frss/publications/2001088/>
- Nguyen, H. 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. doi: 10.1037/a0012702.
- Owston, R. D., Sinclair, M., & Wideman, H. (2008). Blended Learning for Professional Development: An Evaluation of a Program for Middle School Mathematics and Science Teachers. *Teachers College Record, 110*, 1033-1064. 10.1111/j.1467-8535.2008.00875.x
- Peters, W. (1971). *A class divided*. Garden City, NY: Doubleday & Company.
- Ponterotto, J. P., Bulach, S. Greig, T., & Rivera, L. (1998). Development and initial score validation of the Teacher Multicultural Attitude Survey. *Educational and Psychological Measurement, 58*, 1002-1016. doi: 10.1177/0013164498058006009
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*, 1-15. doi:10.1016/S0001-6918(99)00050-5
- Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues, 57*, 55-71. doi: 10.1111/0022-4537.00201
- Riegle-Crumb, C. (2007). Science achievement. In B. J. Bank (Ed.), *Gender and education: An encyclopedia*, (Vol. 1, pp. 449-455). Westport, CT: Praeger.
- Rios, K., Cheng, Z., Totton, R., & Shariff, A. (2015). Negative Stereotypes cause Christians to underperform in and disidentify with science. *Social Psychological and Personality Science, 6*, 959-967. doi: 10.1177/1948550615598378

- Ross, C. E., & Wu, C. L. (1996). Education, age, and the cumulative advantage in health. *Journal of Health and Social Behavior*, 37, 104-120. doi:10.2307/2137234
- Rydell, R. J., McConnell, A. R., & Beilock, S. L. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology*, 96, 949-966. doi: 10.1037/a0014846.
- Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology*, 99, 883-896. doi: 10.1037/a0021139
- Sarason, I. G. (1981). Test anxiety, stress, and social support. *Journal of Personality*, 49, 101-114. doi: 10.1111/j.1467-6494.1981.tb00849.x
- Sarason, I. G., & Palola, E. G. (1960). The relationship of test and general anxiety, difficulty of task, experimental instructions to performance. *Journal of Experimental Psychology*, 59, 185-191. doi: 10.1037/h0047062
- Schimmel, J., Arndt, J., Banko, K. M., & Cook, A. (2004). Not all self-affirmations were created equal: The cognitive and social benefits of affirming the intrinsic (vs. extrinsic) self. *Social Cognition*, 22, 75-99. doi: 10.1521/soco.22.1.75.30984
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38, 194-201. doi: 10.1006/jesp.2001.1500
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115, 336-356. doi: 10.1037/0033-295X.115.2.336.

- Shapiro, J. S., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11, 107-130.
doi: 10.1177/1088868306294790
- Sharma, S. (2005). Multicultural education: Teachers' perceptions and preparation. *Journal of College Teaching and Learning*, 2, 53-63. doi: 10.19030/tlc.v2i5.1825
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: self-affirmation theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 38, pp. 183-242). San Diego, CA: Academic.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28. doi: 10.1006/jesp.1998.1373.
- Spielberger, C. D. (1972). *Anxiety: Current trends in theory and research*. NY: Academic Press.
- Spielberger, C. D., Gorsuch, R. L., Lushene, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302). New York, NY: Academic Press.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629. doi: 10.1037/0003-066X.52.6.613
- Steele, C. M. (2010). *Whistling Vivaldi: Emotion and adaptation*. Oxford, UK: Oxford University Press.

- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 69, 797-811. doi: 10.1037/0022-3514.69.5.797
- Steele, C. M. & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, 45, 5-19. doi: 10.1037/0022-3514.45.1.5
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379 – 440). New York, NY: Academic Press.
- Steele, F. (1975). *Consulting for organizational change*. Amherst, MA: University of Massachusetts Press.
- Steele, S. (1990). *The content of our character: A new vision of race in America*. New York, NY: St. Marten's Press.
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- 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. doi: 10.1016/j.jesp.2007.02.006
- Stricker, L. J. & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, 34, 665-693. doi: 10.1111/j.1559-1816.2004.tb02564.x
- Tabachnick, B.G., & Fidell, L.S. (2001), *Using Multivariate Statistics* (5th ed.). Needham Heights, MA: Pearson.

- Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behavior. *Psychology of intergroup relations*, 7-24.
- Taylor, V. J. & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin*, 37, 1055-1067. doi: 10.1177/0146167211406506
- Thurstone, L.L. (1954). An analytical method for simple structure. *Psychometrika*, 17, 173-182. doi: 10.1007/BF02289182
- Teo, T. & Le Fevre, D. (2016). The development and validation of the Teachers' Perception of Risk Scale (TPRS) in educational change. *Current Psychology*, 36, 649-656. doi: 10.1007/s12144-016-9453-z
- Thompson, B. & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*. 56, 197-208. doi: 10.1177/0013164496056002001
- Thompson, B., & Levitov, J.E. (1985). Using microcomputers to score and evaluate test items. *Collegiate Microcomputer*, 3, 163-168.
- Walton, G. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, 23 (1), 73-82. doi: 10.1177/0963721413512856
- Warner, R. M. (2012). Applied statistics: From bivariate through multivariate techniques. Thousand Oaks, CA: Sage
- Watts, A. L. (2001). *Education and the common good: Social benefits of going to college*. Frankfort, KY: Kentucky Long-term Policy Research Center. 9-11.

- Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology*, 65, 1093-1104.
10.1037/0022-3514.65.6.1093
- Wilson, K. (1997). *Multicultural education*. Retrieved from
<http://www.edchange.org/multicultural/papers/keith.html>
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321-327.
- 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, 459-482. doi: 10.1002/cne.920180503
- Yeung, N. C. J., & von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis & Prevention*, 40, 667-674.
doi: 10.1016/j.aap.2007.09.003.
- Yopyk, D. J. A., & Prentice, D. A. (2005). Am I an athlete or a student? Identity salience and stereotype threat in student-athletes. *Basic and Applied Social Psychology*, 27, 329-336.
doi: 10.1207/s15324834basp2704_5

Appendix A

ST Video Module Transcript

A Mini Course on Stereotype Threat

The purpose of this video is to educate pre-service teachers about a multicultural phenomenon related to the effects of stereotypes. After watching this video you will be able to:

Describe the concept of stereotype threat and its components, recognize stereotype threat in others, and identify and apply appropriate strategies to prevent stereotype threat.

What is stereotype threat? Stereotype threat is the experience of a perceived threat when individuals are concerned about confirming a stereotype about one's group. It is a term that Steele and Aronson developed to explain how stereotypes may have an impact on someone's performance in challenging situations.

How do you know someone is experiencing stereotype threat?

Anyone can experience stereotype threat. Stereotype threat happens when people are in a challenging situation and

- are invested in an area of performance (such as math, writing, or sports)
- are conscious of a stereotype
- and hold a belief that the stereotype could be confirmed as true about their group or themselves through their actions.

For example, Noah, a boy in your class wants to be a basketball player when he grows up. However, he rarely plays basketball. When he does he appears to be very self-conscious and seems to be much more clumsy than usual and sometimes trips over his own feet. One day as he was picking himself up after a fall you hear him mutter "White boys can't play basketball. I'm proof."

Noah appears to be experiencing stereotype threat. He is invested in an area of performance, he wants to be a basketball player.

He is conscious of a stereotype that is relevant to his identity "White boys can't play basketball" And he holds the belief that the stereotype could be confirmed through his actions. It appears that he believes that he is confirming the stereotype through saying White boys can't play basketball. I'm proof.

So why care about it? What is the impact of stereotype threat?

Stereotype threat has been shown to have an impact on children's emotional and academic functioning. For instance, stereotype threat can cause students to score lower on academic tasks and, when experienced over time, can cause students to lose interest in the areas relevant to stereotype threat, like in math or school, in general. Stereotype threat could be one reason why women are under-represented in math related fields. Additionally, losing interest in academics on an individual level can be very costly for a student. Loss of interest in academics can lead to decreased grades and eventual discontinuation of school. For individuals, the lack of a high school diploma is associated with less average lifetime income and the increased likelihood of poor health later in life. Thus, it is important for teachers to know what stereotype threat is and how to prevent students' from experiencing it.

How do you prevent stereotype threat? Now that you know how to recognize stereotype threat as well as the effects that stereotype threat can have, I will present six different ways in which stereotype threat can be prevented.

ST Video Module Transcript Cont.

You can't just want to prevent stereotype threat, you need to D.E.S.I.R.E. to prevent it. D- describe the task differently, E-emphasize another identity, S-self affirmation, I- incremental view of ability, R- role model, and E- external attribution for difficulty.

D-Describe the task differently. One strategy that prevents stereotype threat is changing the way the task is described. For example, a teacher could describe the basketball drills to Noah as an activity to see his ability to be a team player instead of describing it as a basketball related task.

E-Emphasize another identity. Thinking about another identity rather than a threatened identity can prevent stereotype threat. For instance, a teacher may want to remind Noah that he isn't just a White boy; he is also an athlete.

S-Self-affirmation. Instructing people to think about their positive characteristics prevents stereotype threat. For instance, you could encourage Noah to think about his skills or positive qualities before he plays basketball.

I-Incremental view of ability. People who believe that ability can increase through practice could prevent stereotype threat. Noah's teacher could encourage Noah to continue practicing and learning how to be a better player and reassure him that his abilities will grow over time.

R-Role model. Another way of preventing stereotype threat is to provide a role model to the students. In regard to Noah, a teacher may want to show him examples of successful White basketball players, such as Larry Bird and Steve Nash.

E-External attributions for difficulty. Another method that prevents stereotype threat is providing an explanation for why they may be experiencing difficulty. For example, a teacher could explain to Noah that it is normal to be nervous and that he isn't a bad basketball player. He is just experiencing a decrease in performance due to stereotype threat.

Thanks for watching.

Appendix B

ST TOA in Pilot 1 and 2

Please read the following scenario:

Sophia is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is worried about the math unit test. She stated, “Tests in math are hard. I have to do good or else I can’t go on the science field trip and then if I can’t go to that I’ll never be a scientist.” The field trip was too expensive for some students’ families, but, fortunately, your school offered field trip funds to students who were doing well in all their subjects, but could not afford to pay for field trips. Sophia’s parents shared with you earlier that week that she would be unable to attend the field trip unless the school could fund her way.

In the past, you have noticed that Sophia is a bad test taker in math. She works hard in your class and does a nice job on her assignments, but when you grade Sophia's math tests they do not seem to accurately show what she knows. Sophia has strong problem solving skills and is fluent with simple addition, but on tests, addition seems to take her a much longer and she has difficulty determining appropriate strategies to use even though she does not have this difficulty during class assignments. On her most recent math test, she appeared to not have enough time to complete the test, because several questions towards the end were left blank. When you asked Sophia about her most recent math test, she stated “I tried really hard, but my brother says ‘Girls can’t do math.’ I don’t think I got a good grade and now everyone will think girls can’t do math.”

As Sophia’s teacher, what would you put in place to help her achieve what you think she is capable of in math? Please read the following strategies and rate how likely you are to use each strategy to help Sophia.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Take no action.	☐	☐	☐	☐	☐
Explain that it’s normal to feel nervous or anxious on tests.	☐	☐	☐	☐	☐
Have your students take turns giving each other praise.	☐	☐	☐	☐	☐
Emphasize progress instead of a single product.	☐	☐	☐	☐	☐
Remind Sophia that she is strong student.	☐	☐	☐	☐	☐

ST TOA Pilot 1 & 2 Cont.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Reward her whenever she completes her work.	☐	☐	☐	☐	☐
Explain math tests as puzzles using numbers.	☐	☐	☐	☐	☐
Create math activities that are related to her interests.	☐	☐	☐	☐	☐
Allow for Sophia to have a peer model to work with.	☐	☐	☐	☐	☐
Additional tutoring during recess.	☐	☐	☐	☐	☐
Give her additional math worksheets for practice at home.	☐	☐	☐	☐	☐

ST TOA in Actual Study

Please read the following scenario:

Sophia is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is worried about the math unit test. She stated, “Tests in math are hard. I have to do good or else I’ll never be a mathematician.”

In the past, you have noticed that Sophia has not been a good test taker regarding math. She works hard in your class and does a nice job on her assignments, but when you grade Sophia's math tests they do not seem to accurately show what she knows. Sophia has strong problem solving skills and is fluent with simple addition, but on tests, addition seems to take her longer to do. She seems to have difficulty determining appropriate strategies to use even though she does not have this difficulty during class assignments. On her most recent math test, she appeared to not have enough time to complete the test, because several questions towards the end were left blank. When you asked Sophia about her most recent math test, she stated “I tried really hard, but my brother says 'Girls can’t do math.’ I don't think I got a good grade and now everyone will think girls can’t do math.”

As Sophia’s teacher, what would you put in place to help her achieve what you think she is capable of in math?

Please read the following strategies and rate how likely you are to use each strategy to help Sophia.

ST TOA in Study Cont.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Take no action.	☐	☐	☐	☐	☐
Explain that it's normal to feel nervous or anxious on tests.	☐	☐	☐	☐	☐
Have your students take turns giving each other praise.	☐	☐	☐	☐	☐
Emphasize progress instead of a single product.	☐	☐	☐	☐	☐
Remind Sophia that she is strong student.	☐	☐	☐	☐	☐
Reward her whenever she completes her work.	☐	☐	☐	☐	☐
Explain math tests as puzzles using numbers.	☐	☐	☐	☐	☐
Create math activities that are related to her interests.	☐	☐	☐	☐	☐
Allow for Sophia to have a peer model to work with.	☐	☐	☐	☐	☐
Additional tutoring during recess.	☐	☐	☐	☐	☐
Give her additional math worksheets for practice at home.	☐	☐	☐	☐	☐

Appendix C

No-ST TOA in Pilot 1 and 2

Please read the following scenario:

Sophia is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is not concerned about the math unit test. She stated, “I don’t need to do math. I’m going to be famous, and famous people don’t go on science field trips.” Your school offers field trip funds to students who cannot not afford to pay for the field trip and are doing well in all their subjects. Sophia’s parents shared with you earlier that week that she would be unable to attend the field trip unless the school could fund her way. Unfortunately, Sophia may not be able to get the funds needed to attend the field trip due to her poor marks in math and science.

In the past, you have noticed that Sophia is not completing her assignments. Even when she is given extra time in class to complete her work, you need to remind her often to stay on task and to stop talking with her friends. When you grade Sophia's math homework and her tests, few problems are completed. Sometimes she quickly and correctly completes tests, and other times she will turn in a page filled with doodles. Whenever her test is filled with doodles, you always give her a new test and make her complete it. On her most recent math test, she quickly completed her test and when you checked it most of the problems were correct and all were attempted. When you asked Sophia about her most recent math test, she stated, “I wanted to draw. You would have made me do the test again if I drew on the test, so I just did the test fast so I could draw.”

As Sophia’s teacher, what would you put in place to help her achieve what you think she is capable of in math?

Please read the following strategies and rate how likely you are to use each strategy to help Sophia.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Take no action.	☐	☐	☐	☐	☐
Explain that it’s normal to feel nervous or anxious on tests.	☐	☐	☐	☐	☐
Have your students take turns giving each other praise.	☐	☐	☐	☐	☐
Emphasize progress instead of a single product.	☐	☐	☐	☐	☐
Remind Sophia that she is strong student.	☐	☐	☐	☐	☐
Reward her whenever she completes her work.	☐	☐	☐	☐	☐

No-ST TOA in Pilot 1 and 2 Cont.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Explain math tests as puzzles using numbers.	☐	☐	☐	☐	☐
Create math activities that are related to her interests.	☐	☐	☐	☐	☐
Allow for Sophia to have a peer model to work with.	☐	☐	☐	☐	☐
Additional tutoring during recess.	☐	☐	☐	☐	☐
Give her additional math worksheets for practice at home.	☐	☐	☐	☐	☐

No-ST TOA in Actual Study

Please read the following scenario:

Sophia is a student in your second grade class. She is bright and gets along well with her peers. During a transition to math, you overhear Sophia sharing with a friend that she is not concerned about the math unit test. She stated, “I don’t need to do math. I’m going to be famous, and famous people don’t do math.”

In the past, you have noticed that Sophia is not completing her assignments. Even when she is given extra time in class to complete her work, you need to remind her often to stay on task and to stop talking with her friends. When you grade Sophia's math homework and her tests, few problems are completed. Sometimes she quickly and correctly completes tests, and other times she will turn in a page filled with doodles. Whenever her test is filled with doodles, you always give her a new test and make her complete it. On her most recent math test, she quickly completed her test and when you checked it most of the problems were correct and all were attempted. When you asked Sophia about her most recent math test, she stated, “I wanted to draw. You would have made me do the test again if I drew on the test, so I just did the test fast so I could draw.”

As Sophia’s teacher, what would you put in place to help her achieve what you think she is capable of in math?

Please read the following strategies and rate how likely you are to use each strategy to help Sophia.

No-ST TOA in Actual Study Cont.

	Unlikely	Somewhat Unlikely	Undecided	Somewhat Likely	Likely
Take no action.	☐	☐	☐	☐	☐
Explain that it's normal to feel nervous or anxious on tests.	☐	☐	☐	☐	☐
Have your students take turns giving each other praise.	☐	☐	☐	☐	☐
Emphasize progress instead of a single product.	☐	☐	☐	☐	☐
Remind Sophia that she is strong student.	☐	☐	☐	☐	☐
Reward her whenever she completes her work.	☐	☐	☐	☐	☐
Explain math tests as puzzles using numbers.	☐	☐	☐	☐	☐
Create math activities that are related to her interests.	☐	☐	☐	☐	☐
Allow for Sophia to have a peer model to work with.	☐	☐	☐	☐	☐
Additional tutoring during recess.	☐	☐	☐	☐	☐
Give her additional math worksheets for practice at home.	☐	☐	☐	☐	☐

Appendix D

Classroom Organization Measure in Pilot 1 and 2

Please answer the following questions about classroom organization.

	True	False
Good classroom organization ensures that the teacher can see each student at all times.	☐	☐
It is not important for teachers to evaluate if students have easy access to commonly used student tools and resources.	☐	☐
It is not necessary for teachers to move freely and monitor students if the classroom is well organized.	☐	☐
The capability of students to transition from one class activity to the next is closely related to the organization of the classroom.	☐	☐
When organizing a classroom, it is important for teachers to identify the best location to give instruction.	☐	☐
Teachers can encourage or discourage talking among student groups depending on the arrangement of the furniture.	☐	☐
Teachers should increase the number of stimulating visuals present in the classroom.	☐	☐
A well-organized classroom can decrease disruptive behavior and increase academic engagement.	☐	☐

Classroom Organization Measure in the Actual Study

Please answer the following questions about classroom organization.

	True	False
It is not important for teachers to ensure that students have easy access to commonly used student tools and resources (such as textbooks, worksheets, and other classroom materials).	☐	☐
It will be not necessary for teachers to move freely and monitor students if the classroom is well organized.	☐	☐
When organizing a classroom, it is important for teachers to identify the best location(s) to give instruction.	☐	☐
Teachers can encourage or discourage talking among student groups depending on the arrangement of the furniture.	☐	☐
Teachers should have as many stimulating visuals present in the classroom as possible.	☐	☐

Appendix E

TOK ST Questions in Pilot 1 and 2

Please answer the following questions about stereotype threat.

	True	False
Individuals may experience stereotype threat without knowing about their groups' stereotype.	☐	☐
Stereotype threat only affects those who are invested in their performance.	☐	☐
Only females experience stereotype threat.	☐	☐
Students who experience chronic bouts of stereotype threat do not lose interest in the area relevant to the stereotype threat.	☐	☐
Teachers can prevent the effects of stereotype threat by using prevention strategies.	☐	☐
Stereotype threat only affects performance in areas related to school.	☐	☐
Students who believe their actions confirm stereotypes do not experience stereotype threat.	☐	☐
Someone needs to explicitly teach a stereotype in order for a student to experience stereotype threat.	☐	☐

TOK Multicultural Teaching Questions in Pilot 1 and 2

Please answer the following questions about multicultural teaching.

	True	False
Teachers need to be "color blind" and not see the racial or ethnic background of their students.	☐	☐
Sometimes, teachers need to go outside of their comfort zone to question their own beliefs to become more multiculturally sensitive.	☐	☐
To meet the needs of students from different backgrounds, it is important for teachers to understand inequities in the educational system.	☐	☐
Teachers should have positive stereotypes about groups of students, not negative ones.	☐	☐
Teachers should offer an integrated multicultural curriculum on a regular basis, not just during certain months or holidays.	☐	☐
Teachers do not need to present diverse perspectives of the disciplines they are teaching.	☐	☐
A teacher should modify his or her approach to teaching, so students from various backgrounds can access the curriculum.	☐	☐
Teachers will not need to reduce prejudice among their students, if the teachers themselves do not have prejudices.	☐	☐

TOK in Actual Study

Please answer the following questions about instruction, related teaching skills, and situations a teacher may encounter.

	True	False
When addressing negative classroom behaviors, it is important to know what prompts and reinforces the negative problem behavior (BM).	☐	☐
Teachers should have positive views about all groups of students (MC).	☐	☐
When a teacher notices school wide behavior issues, it may be helpful to involve or create a school improvement team (BM).	☐	☐
Students may experience the threat of stereotypes without knowing about their groups' stereotypes (ST).	☐	☐
Facilitating discussion to encourage students to think deeply about a text is an important part of reading instruction (RI).	☐	☐
Sometimes, teachers need to go outside of their comfort zone to question their own beliefs to be more responsive to all students (MC).	☐	☐
When teaching young readers to comprehend text, it is not necessary to teach students to recognize text structure (RI).	☐	☐
Only females experience the threat of stereotypes (ST).	☐	☐
When possible, a teacher should try to encourage parents and guardians to be an active partner in reinforcing appropriate student behaviors (BM).	☐	☐
It is not necessary for teachers to see the cultural background of their students in order to be effective as a teacher (MC).	☐	☐
When selecting teaching materials for reading, teachers do not need to select texts that specifically support their teaching goals (RI).	☐	☐
Stereotype threat only affects performance in areas related to school (ST).	☐	☐
Teaching students reading comprehension strategies like making predictions, visualizing the text, story retelling, and applying prior knowledge are all ways to improve your students' ability to comprehend texts (RI).	☐	☐
Teachers should offer an integrated curriculum on a regular basis, inclusive of cultural issues instead of doing so during certain months or holidays (MC).	☐	☐
Once classroom expectations are established, additional practice, reinforcement, and revisiting of classroom expectations are not needed (BM).	☐	☐
The threat of stereotypes only affects those who are invested in their performance (ST).	☐	☐
When working to decrease a negative student behavior, it is important to reinforce when the student when he or she displays positive replacement behaviors (BM).	☐	☐

TOK in Study Cont.

	True	False
Students who experience the threat of stereotype about a particular area over time may lose interest in that area (ST).	☐	☐
To meet the needs of students from different backgrounds, it is important for teachers to understand the disparities in the educational system (MC).	☐	☐
When teaching reading, it is important to establish an engaging context in which to teach (RI).	☐	☐

Note. MC = Multicultural Teaching, RI = Reading Instruction, ST = Stereotype Threat, and BM = Behavioral Management

Appendix F

Recruitment Email

Dear students,

I know you are all very busy as the semester comes to a close, but if you have some time, you are invited to participate in a survey that examines a professional development learning tool and teachers' ability to use their skills in certain classroom situations. The survey will consist of specific situations and questions regarding those teaching situations. It is expected that it will take you about 20-30 minutes to complete.

As a token for completing the survey, you will be eligible to receive a \$5 gift card for Amazon.com OR two point extra credit.

Your responses will remain confidential and will not be associated with your name or any other information that could otherwise identify you.

Although your participation in the survey is voluntary, I appreciate your cooperation in completing this survey.

If you choose to complete the survey, please do so by **Month XX**, and to access the survey click on the link below:

https://pennstate.qualtrics.com/SE/?SID=SV_0TCRcidTa0m2DuB

Chloe Barry M. Ed.
School Psychology
The Pennsylvania State University

Appendix G

Correlations for TOK and TOA

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. TOK1	--	.16	.05	-.02	.04	-.02	.03	-.10	-.01	-.00	-.05	.10	.02	.20	-.06	-.06
2. TOK2		--	.27	.02	.32	-.15	.01	-.23	-.01	-.13	-.06	-.04	-.18	-.06	-.08	-.11
3. TOK3			--	-.07	.08	-.00	.26	-.05	.05	.06	-.05	-.01	-.21	.05	-.09	-.09
4. TOK4				--	.18	.25	.10	.17	.23	.24	.13	-.03	-.03	.04	-.07	-.08
5. TOK5					--	-.16	-.20	-.33	-.13	-.29	-.10	-.23	.02	.12	-.12	-.11
6. TOA1						--	-.07	.11	-.17	.19	-.12	-.03	-.05	-.24	.13	.01
7. STP1							--	.45	.04	.44	.13	.17	-.20	.00	.04	.31
8. STP2								--	.33	.32	.16	.06	.08	.22	.30	.36
9. STP3									--	.07	.07	.03	.25	.51	.12	.07
10. GP1										--	.28	.18	-.04	.00	.13	.13
11. GP2											--	.19	.06	.29	.17	.32
12. GP3												--	.04	.09	.21	.20
13. ISTP1													--	.49	.06	-.06
14. ISTP2														--	-.04	.00
15. ISTP3															--	.62
16. ISTP4																--

Note. $N = 60$; 1 = TOK: Class Org Total; 2 TOK: Behavior Total; 3 = TOK: Read Total; 4 = TOK: ST Total; 5 = TOK: MC Total; 6 = TOA: Take No Action; 7 = ST Prevention: Explain feeling anxious is normal; 8 = ST Prevention: Explain math tests are puzzles using numbers; 9 = ST Prevention: Remind her that she is a strong student; 10 = Generalized Prevention: Students take turns giving each other praise; 11 = Generalized Prevention: Allow a peer model to work with her; 12 = Generalized Prevention: Emphasize progress over a single product; 13 = Ineffective ST Prevention: Give rewards for completed work; 14 = Ineffective ST Prevention: Create math activities related to her interests; 15 = Ineffective ST Prevention: Additional tutoring during recess; 16 = Ineffective ST Prevention: Give additional worksheets to practice at home; Means for measures of strategies were based on the Likert scale ranging from 1 (*unlikely*) to 5 (*likely*) and ST measures were based on the Likert rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Appendix H

Summary of EFA, Maximum Likelihood Extraction, Promax Rotation of TOA

Item	Factor I	Factor II	h^2
TOK-II 11 - Give worksheets to practice at home	.82	-.10	.67
TOK-II 10 – Additional tutoring during recess	.68	-.12	.46
TOK-II 9 – Allow a peer model to work with her	.51	.16	.21
TOK-II 7 – Explain math tests are puzzles using numbers	.40	-.05	.30
TOK-II 2 – Explain feeling anxious is normal	.36	.25	.16
TOK-II 3 – Students take turns giving each other praise	.31	-.04	.09
TOK-II 4 – Emphasize progress over a single product	.27	.05	.08
TOK-II 8 – Create math activities related to her interests	-.01	1.00	1.00
TOK-II 6 – Give rewards for completed work	-.07	.52	.25
TOK-II 5 – Remind her that she is a strong student	.13	.49	.28
TOK-II 1 – Take no action	.07	-.24	.06
Eigenvalues	2.64	1.96	
% Variance accounted for	24.00	17.84	
Cronbach's α	.68	.63	

Note. $N = 70$.

Factor I = Reactive Strategies; Factor II = Proactive Strategies; Salient ($\geq |.40|$) pattern coefficients are in bold.

Appendix I

Frequency Table of Response Rates for TOK and Classroom Organization ($N = 60$)

Stereotype Threat	True		False	
	Frequency	Percentage	Frequency	Percentage
Item 1	59	98.30%	1	1.70%
Item 2	0	0.00%	60	100.00%
Item 3	0	0.00%	60	100.00%
Item 4	3	5.00%	57	95.00%
Item 5	58	96.70%	2	3.30%
Multicultural Teaching				
Item 1	58	96.70%	2	3.30%
Item 2	58	96.70%	2	3.30%
Item 3	3	5.00%	57	95.00%
Item 4	57	95.00%	3	5.00%
Item 5	60	100.00%	0	0.00%
Reading Instruction				
Item 1	60	100.00%	0	0.00%
Item 2	7	11.70%	53	88.30%
Item 3	13	21.70%	47	78.30%
Item 4	58	96.70%	2	3.30%
Item 5	59	98.30%	1	1.70%
Behavior Management				
Item 1	58	96.70%	2	3.30%
Item 2	59	98.30%	1	1.70%
Item 3	59	98.30%	1	1.70%
Item 4	4	6.70%	56	93.30%
Item 5	58	96.70%	2	3.30%
Classroom Organization				
Item 1	9	15.00%	51	85.00%
Item 2	0	0.00%	60	100.0%
Item 3	60	100.00%	0	0.00%
Item 4	59	98.30%	1	1.70%
Item 5	32	53.30%	28	46.70%

VITA
Chloe Y. H. Barry

EDUCATION

Doctoral Candidate, School Psychology, The Pennsylvania State University	Anticipated Grad. 2017
M.Ed., School Psychology, The Pennsylvania State University	May 2012
B.A., Psychology, <i>summa cum laude</i> , Arizona State University	May 2009

CERTIFICATES

School Psychologist , Pre-Kindergarten to Grade 12, Florida	2016
Educational Specialist II , School Psychologist, Pennsylvania	2015
School Psychologist , Administration and Pupil Personnel Services, Provisional Certificate, New York	2015
School Psychologist , Educational Services Personnel, New Jersey	2015
School Psychologist , Special Services, Connecticut	2015

TEACHING EXPERIENCE

Teaching Assistant , The Pennsylvania State University	2014-2015
Teaching Assistant , The Pennsylvania State University	2010-2012
Tutor , Bellefonte Elementary School, Bellefonte, PA	2009
Teacher , KinderCare Learning Center, Kent, WA	2005-2007

CLINICAL EXPERIENCE

Post Doctoral Fellow , University of Central Florida	2017
School Psychologist , Broward County School District, Fort Lauderdale, FL	2016-2017
School Psychologist , New York City Department of Education, New York City, NY	2016
School Psychologist Intern , Montgomery County School District, Silver Spring, MD	2013-2014
School Psychologist Substitute , State College Area School District, State College, PA	2012-2013
School Psychology Intern , State College Area School District, State College, PA	2011-2012
Student Clinician , The Pennsylvania State University	2010-2012
Student Clinician , Service Learning, Arizona State University	2008

PUBLICATIONS

Barry, C. Y. H., & Okun, M. (2011). Using investment theory to predict decay of the intent to persist among freshmen. *Journal of College Student Retention*.13(1), 87-107. doi:10.2190/CS.13.1.e

PROFESSIONAL PRESENTATIONS

Barry, C. Y. H., & Vandiver, B. J. (2011, August). *The effects of time perspective and ethnic identity on Asian and Caucasian American secondary level students' perceived academic achievement*. Poster session to be presented at the annual American Psychological Association Conference, Washington D. C.

Barry, C. Y. H., & Okun, M. (2009, April). *Using investment theory to predict decay of the intent to persist among freshmen*. Poster session presented at the annual Celebrating Honors Symposium of Research and Creative Projects, Arizona State University, Tempe, AZ.