


12-2010

Verbal cues: Producing the same results in stereotype threat research?

Tarryn E. McGhie

University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>

 Part of the [Developmental Psychology Commons](#), [Educational Psychology Commons](#), and the [Science and Mathematics Education Commons](#)

Repository Citation

McGhie, Tarryn E., "Verbal cues: Producing the same results in stereotype threat research?" (2010). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 756.

<https://digitalscholarship.unlv.edu/thesesdissertations/756>

This Thesis is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

VERBAL CUES: PRODUCING THE SAME RESULTS
IN STEREOTYPE THREAT RESEARCH?

by

Tarryn Emeka McGhie

Bachelor of Arts
North Carolina Agricultural and Technical State University
2007

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Educational Psychology
Department of Educational Psychology
College of Education

Graduate College
University of Nevada, Las Vegas
December 2010

Copyright by Tarryn Emeka McGhie 2011
All Rights Reserved



THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Tarryn Emeka McGhie

entitled

Verbal Cues: Producing the Same Results in Stereotype Threat Research?

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Educational Psychology

Rebecca Nathanson, Committee Chair

Gita Taasoobshirazi, Committee Member

Gregory Schraw, Committee Member

William Cross, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

December 2010

ABSTRACT

Verbal Cues: Producing the Same Results in Stereotype Threat Research?

by

Tarryn Emeka McGhie

Dr. Rebecca Nathanson, Examination Committee Chair
Professor of Education and Law
University of Nevada, Las Vegas

This study examined the notion that stereotype threat experiments can be influenced through linguistic manipulation. The cueing of a phrase (whether stereotypical or non-stereotypical) can produce performance differences between groups, rather than cueing of a stereotype, as used in previous research. Participants ($n=95$) mostly Caucasian females (68%) ranging in age from 18-45 ($M=22.7$). The design involved three groups and participants were randomly assigned in order to control for consequential affects. The control group received no verbal cues. The stereotypical group received a stereotypical cue (i.e. men tend to do better on this test than women). The counter-stereotypical group received a false stereotypical cue (i.e. women tend to do better on this test than men do). After cueing, all participants completed a math test. The General Record Examination (GRE) was used and the dependent measure was the participant's score on the test. The results of an F test show there was no interaction between the group assignment (assigned cue) and test score in relation to number of items attempted ($F_{(2, 94)} = .968, p>.05$), correct ($F_{(2, 94)} = .193, p>.05$) and the difference between correct/attempted ($F_{(2,94)}=1.450, p>.05$).

TABLE OF CONTENTS

ABSTRACT	iii
CHAPTER 1 INTRODUCTION	1
Purpose and Significance	2
Research Questions	4
Definition of Terms	4
CHAPTER 2 REVIEW OF RELATED LITERATURE	6
General Background	6
Verbal Cues	7
Gender Differences	9
Building the Hypothesis	11
CHAPTER 3 METHODOLOGY	13
Participants	13
Measures	13
Procedure	14
Post Test Completion	15
Treatment of Data	16
CHAPTER 4 RESULTS	17
Statistical Analysis of Research Questions	17
Test Performance	18
CHAPTER 5 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	22
Discussion of Results	22
Limitation	22
Recommendations	23
Conclusions	24
REFERENCES	25
APPENDIX	29
A. Survey	
B. Demographic Sheet	
C. Subject Pool Ad	
D. Debriefing	
E. Tables 1, 2, and 3	
VITA	38

CHAPTER 1

INTRODUCTION

The theories on stereotype threat first introduced by Claude Steele gives some explanation as to why certain groups perform lower on specific tasks. Steele and Aronson (1995) define stereotype threat as “being at risk of confirming, as self-characteristic, a negative stereotype about one’s group” (Davis, Aronson, & Salinas, 2006). Stereotype threat has been tested to show that Whites will perform lower on athletic tasks when compared to Blacks, Whites will score lower on intelligence tests when compared to Asians, women will perform lower on math and science tasks when compared to men, Blacks will score lower on intelligence tests when compared to Whites, and Latinos will score lower on intelligence tests in comparison to all other racial groups (Steele, 1997). The researcher on stereotype threat states that individuals will perform poorly when the stereotype is primed during a situation in which performance is measured (Jamieson & Harkins, 2007). Stereotype threat mechanisms have yet to be isolated, but it is theorized to be linked to anxiety, expectancy, arousal, working memory interference, cognitive load, reactance, and withdrawal of effort (Jamieson & Harkins, 2007). Though these are credible reasons, some researchers suggest that the reason for the differences in testing outcomes are much more related to socioeconomic status and education; however, these resources may very well be a result of the affects of stereotype threat on a previous generation. This implies that these values of false self evaluation could be past down throughout a family’s history. The current research is guided by the hypothesis that the stereotype threat, if indeed true, is evident in testing situations; therefore there is no need to trigger the stereotype. The research intends to demonstrate

that the current procedure for soliciting stereotypical threat is not an activation of an actual stereotype, but it is in fact the result of a verbal cue. In other words, the intent of this research is to show that the verbal cue is what causes the difference in groups when testing stereotype threat. The Verbal Cue causes the difference in the test taker, and does so whether the verbal cue is stereotypical or counter-stereotypical. In simple terms, any mention of a stereotype increases the saliency of that stereotype whether or not it is an actual fact.

Purpose and Significance

The purpose of this study is to identify additional significant factors that contribute to the theory of stereotype threat. The intent is to determine whether verbal cues influence the results of test takers. These factors may be the cause, but are not tested directly in this study. These factors may include memory: interference, increase in anxiety, decrease in anxiety, influence on motivational factors for test takers, and a desire for representation of ones own gender through self evaluation of ability. Throughout the years, many researchers have published results that enhance the original thoughts on stereotype threat theory as first introduced by Steele and Aronson in 1995. Though these researchers have replicated and enhanced the study on stereotype threat research, few have challenged the inevitable outside influence that can contribute to the results of testing the theory. Some studies have used Steele's "building block" thoughts on stereotype threat theory that showed a difference between testing for ability and testing for non-diagnostic ability. Other researchers have used part of the original Steele experiment to conduct related studies, such as implementing the collection of biographical data during post test conditions instead of pretest. Although previous

research has made tremendous strides towards understanding this phenomenon, few have challenged the undermining of the theory, nor have they implemented any testing on all of the predetermined limitations of the situational design. For instance, limitations that have been brought to light include, but are not limited to: 1) race of the administrator 2) sex of the administrator 3) gender and race of examinees and 4) pretest instructions.

In long term goals, this study is the foundation I will eventually use to broaden the thought process of researchers when determining the administration of examinations. In addition, participants' background and experience may influence the reliability and validity in similar stereotype threat testing situations. The thought behind this is that if you are "uncomfortable" during a time in which your performance is being measured, you will endure additional anxiety and external needs during the task; furthermore influencing the results, undoubtedly in a declining fashion. For instance, if a participant was experiencing a phobia of spiders (arachnophobia) and was asked to take a test that was not a determination of ability (even if it was) and on the way to the testing area several pictures of spiders were in the hallway leading to the testing area, the participant may be influenced. Although some may say that the influence of this factor may be minute, it can have a significant impact on results of testing regardless of perception during the task. Many researchers down play these factors that can influence the participant; moreover, these factors can be controlled for.

In addition, this study is intended to lay the ground work for similar studies that may be used to enhance standardized test scores. This research intends to show evidence that supports the main premise behind stereotype threat, but offers a remedy for these threats.

Research Questions

There has been much confusion on the issue of saliency in stereotype threat testing. Many studies have shown a difference between the control group and the stereotype threat salient group; on the other hand many other studies have shown that the two groups were identical. Both of these results have reinforced the stereotype threat theory; however, few have used a counter-stereotypical salient group as a resource for testing stereotype threat phenomenon. This research study intends to test whether the opposite of stereotype threat salience can A) reverse the thoughts of the examinee and increase the test score of female test takers in mathematics when relating to their male counterparts or B) presenting the group with a counter-stereotypical cue create salience of the stereotype. As the title describes the main research question of this study is “*Can verbal cues, regardless of factual evidence create stereotype threat salience (telling females that they will do better than males on this exam)?* If the results show a significant difference (i.e. an increase in females test score in the counter-stereotypical group when compared to the test scores of the participants in the stereotypical group and/or the control group), this study will help confirm a resource for combating stereotype threat. If there is no difference between groups, then this will reinforce my premise for conducting this study: regardless of the cue, any verbal cue will trigger stereotype threat and the presence of a test makes it salient within the test taker; therefore, cueing is not needed to induce the threat.

Definition of Terms

Stereotype threat theory - A phenomenon, based on the studies of minorities, originally discovered by Claude Steele and Joshua Aronson in 1995 on the campus of the

University of Stanford in San Francisco, California. It states that certain groups are affected by the salience of a stereotype when in competition with another group. It does not judge the validity of the stereotype in question.

CHAPTER 2

THEORETICAL FRAMEWORK

General Background

The Processing Efficiency Theory implies that triggering and elevating cognitive aptitude is impacted as stress or task anxiety is increased in an individual. Researchers suggest this stress and anxiety can be brought on by stereotype threat (Osborne, 2007). The effects of stress anxiety (i.e. stereotype threat) are more evident as the task difficulty increases. The research on stereotype threat shows that if the task is simple, there are no significant differences between participants; however, the research also shows as the task increases the overwhelming need to defy the stereotype threat also increases. This impairs cognitive ability to succeed at the task, due to occupied space in the working memory by thoughts of the negative stereotype (Bonnot & Croizet, 2007; Osborne, 2007; Jamieson & Harkins, 2007; Stangor, Carr, & Kiang, 1998; & Steele, 1997).

Osborne's (2007) study examined the difference in body temperature, blood pressure and heart rate of males and females in low and high stereotype threat conditions, while taking a mathematics exam. The results of the study showed that females in the high stereotype threat condition had significantly higher body temperature than males and females in all other low and high stereotype threat conditions. The increase in body temperature has been linked to increases in anxiety, stress and arousal in test takers (Osborne, 2007).

Working memory is explained as the continuous cognitive representation of stimulus even after the stimulus has been removed from visual or auditory perception. Often times when discussing working memory, we are discussing multitasking on two or

more mental processes. The difficulty in multitasking becomes evident during working memory load, usually as the specific task difficulty increases. The question becomes interesting as we learn what tasks can be paired with each other. The load theory states that there are increases in interference during active control such as in working memory. Other research that measures visual searching and how it is affected by working memory shows that visual searches were impaired when executive control of the working memory is overloaded. These findings in research suggest that processing information is done by several resources, each with its own threshold (Park, Kim, & Chun, 2007). This suggests that stereotype threat is linked directly with effects of working memory impairment through overload. The current research paper does not measure the effects of working memory load, but the literature provides information about the characteristics of stereotype threat theory and how it is influenced by cognitive processes.

Verbal Cues

The persistence of researchers to explain the effects of stereotype threat has been an up hill battle. Researchers have found that in cases where age, race, education, income and marital status are controlled, performance differences are still present among participants, especially gender differences between males and females. They attribute these differences to stereotype threat and they are continually observed in stereotype threat experiments (McGlone, Aronson, & Kobrynowicz, 2006). Inzlicht and Ben-Zeev (2000) demonstrated that female test takers exhibited significantly better test scores when the testing site was composed of women as the majority in the class room setting. The complete opposite was the case when females were the minority in the test setting. The females in the minority setting scored significantly lower than males as well as their

female counterparts in the majority setting (McGlone, Aronson, & Kobrynowicz, 2006). This study provides evidence that stereotype threat can be influenced by an individual's response to the immediate context. In itself, having males present is a form of creating saliency. The current research paper uses all female test takers in an effort to control for additional anxiety brought on by male "competitors" in the testing environment.

In regards to the effects of the administrator, Davis and Silver (2003) reported that their research provides evidence that the race of the interviewer or researcher of a study made significant differences in responses. They tested Black Americans via telephone interview. This study used Black interviewers and White interviewers. The research confirmed that participants performed significantly better with the Black interviewer than they did with the White interviewer. The researchers attributed the results to added pressure on participants who were interviewed by Whites to negate the stereotype, to show they could perform as well as Whites on the task (McGlone, Aronson, & Kobrynowicz, 2006). The Davis and Silver research is significant to this research. The gender of the administrator may be a factor in affecting test takers stereotype threat salience in mathematics. The impact of this may be minimal for this particular study because no interviews are being conducted; however a task is being measured by and administrator of the opposite sex. This variable is not measured, but may be significant in future research in this area.

While examining other groups subjected to stereotype threat, Steele and Aronson's (1995) study provided evidence that African American students that were asked their race on a demographic sheet prior to testing performed lower on an examination given, when compared to African American participants who were asked to

specify their race after the examination was given (Stangor, Carr, & Kiang, 1998). These results show that very little effort is needed to trigger stereotype threat salience. Stangor et al. go on to conclude that the expectation to succeed on a particular task is jointly influenced by the participant's ability to perform the task (Stangor, Carr, & Kiang, 1998). Stangor et al. also concluded the influence of activating stereotypes supersedes and is stronger than the expectations of the participant's abilities to complete the task (Stangor, Carr, & Kiang, 1998). Stangor et al. research also confirms the recent research regarding the theories of stereotype threat in that indicating race and gender prior to testing is enough manipulation to trigger thoughts of a stereotype resulting in saliency. This research will be used to minimize stereotype threat salience during administration of the task. The identification of biographical information will be collected during post test procedures in order to minimize any outside interference in relation to stereotype threat salience.

General Differences

When studying stereotype threat, specifically its effects on women, there are undeniable gender differences in many cases. These gender differences seem to be prevalent in most testing situations, and are skewed to one gender or the other. In many cases, research is conducted in order to determine why. Recently, researchers have focused their work on stereotype threat research in regards to gender; the most common and well known stereotype threats, in terms of gender are in mathematics and science. Bonnot and Croizet (2007) took an aggressive approach and hoped that a new angle could be used to conduct and review this research. The researchers wanted to look at the self concept and abilities of females with counter-stereotypical majors, those in math and

science. The researchers believed that the females that have decided to enter into fields that are counter-stereotypical, have not internalized the stereotype and therefore were not affected by stereotype threat. The researchers quickly discovered that this was not the case. The results show that female participants in the study that majored in math and science were, for the most part, aware of the gender stereotypes and it did not keep them from this field of interest; however, the female participants still had a low self perception of their own math and science ability. An interesting outcome from the research also showed an opposite effect on males in the study. It showed males as having an increase in their perceived math and science ability. This is due to the “known” gender differences related to mathematics and science (Bonnot & Croizet, 2007). This is significant to this research paper because it shows that self determination is not a deterrent to a career and that females in these counter-stereotypical majors believe that effort will prevail in their circumstances.

Nosek, Banaji, and Greenwald (2002) and Schmader, Johns, and Barquissau (2004) found in their studies that females who enrolled in math and science related fields, continued to view math and science as masculine fields of study. According to stereotype threat theory, the group in most danger of susceptibility to stereotype threat is at risk of internalizing the negative stereotype, especially when the disadvantages are made salient during performance checks (e.g. tests, performance tasks, etc.).

In some recent studies it has been questioned whether some group members associated with a certain stereotype not only show bias of the stereotype, but also may give into the stigma of the stereotype and believe it themselves (Bonnot & Croizet, 2007).

A study by the Organization for Economic Co-operation and Development, Program for International Student Assessment (OECDPISA) reported large differences between males and females in attitudes towards mathematics (Bonnot & Croizet, 2007). Bonnot and Croizet (2007) explain that the more a person endorses the negative stereotype, the lower they will perform on the task at hand; specifically with women in regards to performance in math ability when compared to males. The researcher goes on to explain that the stereotype threat can lead to low self-evaluation and can interfere with performance by triggering a disruptive mental load. This is evident in their study; in which participants are asked to self evaluate themselves on a likert scale, with questions such as “I am bad at math”. Participants are more likely to internalize the negativity of this question and become victims of the self-fulfilling prophecy. The study goes on to state that underachievement of females in this study may be due to the self evaluation of skills may shape an over expectation of success (Bonnot & Croizet, 2007).

Building the Hypothesis

There has been much debate on the differences between how much internal vs. external sources contribute to stereotype threat. Although most of the researchers agree that stereotype threat is an internal acknowledgement of the negative stereotype, the majority of testing has focused on identifying the external sources (Schmader, Johns, & Barquissau, 2004). Even though we can possibly control for many of the external sources (in testing situations), it still can and does contribute to making the stereotype salient, thus contributing to the internalizing effects suffered by test takers in groups that are affiliated with susceptibility.

The previous research presented will be used to support the results of this study. Stereotype threat may be salient; however it is the administrator's duty to create a calm and welcoming environment that encourages ability and effort while controlling for external threats to the test takers mental state. These factors will improve the test takers physical, emotional and mental health, while resulting in an accurate gauge of ability. The factors that this study controls for include: 1) All female test takers, to ensure no additional anxiety is brought on by competition with male counterparts. 2) Biographical information is collected post test to decrease the likelihood of salience. 3) An absence of clocks to ensure that time limits are not a factor in increasing anxiety. 4) students are not permitted to leave until all test takers have completed the biographical information unless they would like their information withdrawn from the study in order to decrease any pressure that may be brought on by early completion of the task.

CHAPTER 3

METHODOLOGY

Participants

There were 95 female students (68.8% White, 13.5% Asian, 11.5% Black, and 5.2% Latino) from the University of Nevada, Las Vegas who participated in the study. The participants' ages ranged between 18 and 45 years of age ($M=22.7$, $SD=4.2$, with 1 participant's information missing). All the participants were undergraduate students, mostly juniors (2 freshmen, 33 sophomores, 54 juniors, 6 seniors). The majority of the participants had parents with some form of college experience. The participants' parental education levels were recognized as the "highest level achieved by either parent", and the results are the following: 20 with graduate degrees, 14 with bachelor degrees, 40 with some college/associate degree, 18 with a high school diploma and 2 without a high school diploma. Fifty-four participants were participating as part of the department subject pool and received research credit in exchange for participation in the study. Forty-one participants were volunteers from outside of the departmental subject pool and received no compensation for participating.

Measures

The measurement outcome used in this study was based on mathematics performance. Participants took a very short version of the General Record Examination (GRE). The reason these particular questions were chosen was based on: A) the original research which used a short version of a proven reliable standardized test most like that of this research; B) These items ranged in difficulty; C) Mathematics is a proven genre for stereotype threat in relation to gender differences; and D) The questions also give the

participant a chance to build confidence in their math ability, which may result in an increase of effort to complete all problems. The examination consisted of 20 quantitative problems which vary in range of difficulty. All the problems were multiple choice. For example:

If the average (arithmetic mean) of $3a$ and $4b$ is less than 50, and a is twice b , what is the largest possible integer value of a ?

(A) 9 (B) 10 (C) 11 (**D) 19** (E) 20

If $3x-4=11$, what is the value of $(3x-4)^2$?

(A) 22 (B) 36 (C) 116 (**D) 121** (E) 256

Procedure

Once the consent form was signed, all participants were asked to take the short version of the GRE. The room was completely quiet and free of distractions. All participants were given a yellow folder, which contained biographical questions. On top of the envelope was the test, and on top of that was a blank sheet of paper that participants could use for scrap (to calculate problems, because a calculator was not allowed). Participants were given 20 minutes to complete the exam. Participants were not informed that the test was being timed. This information was withheld so that no additional anxiety would be brought on by participants who were worried about the time left. Participants were not allowed to proceed to the next step until the 20 minutes lapsed and all participants were notified when time was up. This was to ensure that no additional anxiety was brought on by early completions of the test. All group procedures are the same up to this point. The groups received different instructions from this point

of the study. The instructions are detailed below according to the condition the participants were assigned to.

Control group

The control group was given the exam with no additional “Verbal Cues”. No verbal communication was given to this group other than test instructions detailed above.

Stereotypical Cue group

The SC group was given the exam with a verbal cue that stated the following “Men tend to do better on this test than women do”. This was a verbal cue that is a well known stereotype. There were no additional verbal cues given to this group other than test instructions.

Counter-Stereotypical Cue group

The CSC group was given the exam with a verbal cue that states the following “Women tend to do better on this test than Men do”. This was a verbal cue, but is not a known stereotype. This opposite cue made the group the counter-stereotypical group.

Post Test Completion

After the participants completed the test, they were then asked to complete a demographics sheet to ensure that their gender identification was not a distracter during the test. This also ensured no additional saliency of stereotype threat was brought on by gender identification prior to the task. Once all the tasks were completed, all participants were debriefed on the nature of the study. They were also told the intent of the study and how this information will be used to enhance their testing experiences in the future. The participants were also given the opportunity to have their information removed from the study as well as options to receive the results once the testing and analysis are completed.

No participants asked for their information to be removed from the study. Participants were then asked to keep all the information of the study confidential so that future participants will not be exposed to the information being requested and in turn will not compromise the study.

Treatment of Data

The data from this study were scored as correct or incorrect, due to the test being multiple choice. Although the test consisted of twenty items, the items were scored based on questions attempted and questions correct. Therefore, percentages were calculated by dividing the questions correct by the questions attempted. The test did not have any items with multiple answers. Participants were allowed to skip items if they so desired. A skipped item did not count towards an attempted item. The data for this study, which includes the consent form and the actual test, have been separated into four stacks and inserted into labeled folders: Control group, Counter-stereotypical Group, Stereotypical Group and Consent Forms. All of the information is kept in the locked office of Tarryn McGhie in room 227A of the Carlson Education Building on the campus of the University of Nevada, Las Vegas in Las Vegas, Nevada. Once the data was collected, it was immediately scored and recorded on a laptop with password protection. All of the data collected in this study has been sequestered and only Tarryn McGhie has had access to the information. The information will be stored for three years, as required by the Protection of Human Subjects division of the University of Nevada, Las Vegas.

CHAPTER 4

RESULTS

This chapter details the statistical analysis and the results of the study. The chapter begins with the analysis conducted and then moves to the results of that analysis. The analysis includes the mean, and standard deviation of all dependent and independent variables as well as the results of the analysis of variance (ANOVA). The chapter includes the details of how the analysis was conducted and the results of each in relation to the hypothesis as well as additional analysis results.

Statistical Analysis

In order to identify differences in group variability, an analysis of variance (ANOVA) was performed for a between-groups design. The analysis is intended to identify differences between the conditions of each group. The purpose of the hypothesis is to determine the affect of the counter-stereotypical cues on the performance of participants in the CSC group, in comparison to all other group's performance.

A frequency table was used to identify common frequencies between major of study, age, parental education and number of correct answers on the GRE. This analysis will identify if certain majors with extensive mathematical courses improves or decreases the likelihood of achieving a higher score. Special factors may elevate or diminish achievement on this exam. These include: Parents perceived math ability and education, developmental environment (the home and personal social setting), as well as past experiences in problem solving. These questions will be addressed in the biographical portion of the test administration. It is expected that these participants may be outliers, and their data may be used to make inferences about their particular abilities.

Test Performance

The 20-item GRE practice examination was distributed to the participants (n=95). The overall average correct ($\underline{M}=7.35$) to attempted ratio ($\underline{M}=13.77$) was calculated and yielded a mean score of 55.21. These statistics are shown in table 1.

Hypothesis: CSC group difference \geq SC and CG

An analysis of variance was conducted, and the results showed no significant difference between groups. The first variable described is score percentage. Percentage was calculated by questions correct/questions attempted. The Control group yielded a mean score of 59.2 (SD=22.4) questions correct for every 12.8 (SD=4.7) questions attempted. The Counter-stereotypical cue group yielded a mean score of 49.5 (SD=24.1) questions correct for every 14.4 (SD=4.4) questions attempted}. The Stereotypical cue group yielded a mean score of 56.6 (SD=29.8) questions correct for every 14.1 (SD=5.1).

F-test

An analysis of the number correct, number attempted and the ratio between them was conducted to compare group interactions. The number correct was not statistically significant ($F_{(2, 94)} = .193, p > .05$). There was no interaction between the type of group participants were assigned to and the performance on the examination in relation to number of questions correct. The number attempted was not statistically significant ($F_{(2, 94)} = .968, p > .05$). There was no interaction between the type of group participants were assigned to and the performance on the examination in relation to number of questions attempted. The ratio between the numbers correct and attempted also was not statistically significant ($F_{(2, 94)} = 1.450, p > .05$). There was no interaction between the type of group participants were assigned to and the performance on the examination in relation to the

difference between the questions attempted and the questions correct. The results show that the verbal cues assigned to each group did not alter test performance.

Table 1. Group means and standard deviations

	N	m	SD
Control			
	31	59.2	22.4
Stereotypical			
	34	56.6	29.8
Counter-stereotypical			
	30	49.5	29.1
Total	95	55.2	25.9

Table 2. Group questions attempted and questions correct

	N	m	SD
Control	31		
<i>attempted</i>		12.8	4.7
<i>correct</i>		7.3	3
Stereotypical	34		
<i>attempted</i>		14.1	5.1
<i>correct</i>		7.6	3.8
Counter-stereotypical	30		
<i>attempted</i>		14.4	4.4
<i>correct</i>		7.1	4

Total	95		
<i>attempted</i>		13.7	4.7
<i>correct</i>		7.4	7.3

Correlation Analysis

A Pearson r was conducted to compare perceived level of test difficulty with items correct and items attempted. The participants in the Stereotypical group perceived the study as more difficult than the CSC and control group with an $r = .20$ (sig. at .05 level). A Pearson r was also calculated on all groups, and compared exam difficulty and percentage score. Although all groups showed a weak to moderate negative correlation, the stereotypical group was the only group that was significant at the .05 level ($r = -.425$). (CSC group $r = -.333$, control group $r = -.139$). This analysis is detailed in Table 4.

Table 3. Group means for attempted, correct and the difference between attempted/correct.

	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Level of Significance</i>
Correct					
Between Groups	5.137	2	2.568	.193	.825
Within Groups	1226.4	92	13.350		
Total Attempted	1231.537	94			
Between Groups	44.131	2	22.066	.968	.384
Within Groups	2096.774	92	22.791		
Total Difference	2140.905	94			
Between Groups	46.798	2	23.399	1.450	.240
Within Groups	1484.360	92	16.134		
Total	1531.158	94			

CHAPTER 5

DISCUSSION AND CONCLUSION

Discussion of Results

This study examined the implications of pre test cues and their effect on math performance in women. The study also attempted to identify additional factors that may serve as external threats to validity and reliability when testing stereotype threat and other related constructs. The study predicted that woman in all groups would perform the same due to stereotype threat salience (in reference to mathematics) regardless of the cue or that women who received the counter-stereotypical cue would perform better than all groups due to elevated confidence.

The main inference this research intended to provide is *how this information can be used in standardized testing?* The intent was to apply the principles of the research to standardized testing and other “real world” situations; however no affect was found. This idea was developed because researchers have often over predicted the difference for the minority group and underestimated the *elevating* effects on the majority group in testing (i.e. Whites in intelligence testing when compared to other non-Asian minorities; Males in comparison to females during mathematics and science tasks etc) although this study only tested females.

Limitations

The primary limitation of this study was the use of the subject pool to solicit participants. The design of the study intended to have groups of 10 or more participants in each group during the task; however, the groups ranged from 3-8 participants in each individual testing situation. This was not disclosed to me until after the first testing

sessions was completed. This may have compromised the studies original goal.

Although all participants were female, the group size may not have been large enough to elicit the effect that was originally desired. The nature of the subject pools requirement to provide examples of test questions may also have compromised the participants. This may have created stereotype salience or anxiety prior to their scheduled session.

Recommendations

The research conducted has the theoretical framework to have a significant difference if the participants have more motivation to put forth effort at the task. For instance, if the participants were taking this test for a grade as opposed to research credit, the likelihood that the scores would reflect this would be tremendous. This of course would have to be conducted as a pop quiz or an exam used for students to “test out” of a class. This would be needed because prior exposure to the type of questions would increase the likelihood of success.

If this study was to be conducted again, I would recommend that race be examined instead of gender. Race is more of a contextual trigger to stereotype threat than gender.

Additional Recommendations of Measurement

In addition to previous recommendations for future research, I would recommend an additional measure of performance based on gender. A gender identity measure would have been able to provide additional information on constructs linked to gender identity, gender role and performance in a stereotype threat situation. A measure such as the Conformity to Feminine Norms Inventory (CFNI) could have proved to be a valuable tool of measurement, especially in the correlation analysis.

Conclusions

This study intended to provide additional information to be used for combating stereotype threat in testing situations. This was attempted through following normal stereotype threat guidelines for previous studies, as well as combining the limitations of previous research studies done in the same group and subject matter, females, and math respectively. By combining all limitations in an attempt to control for them, this study lacked balance in other areas including random assignment and sample size of test groups, already discussed in the limitation sections. This study suggests that females may still be negatively affected in test situations, simply due to subject matter; however these conclusions may be caused by other factors previously described in the limitations section. This is likely the case due to a lack of differences in the interactions between test scores across the three group assignments. More research in gender norms, environmental triggers and pre-test performance is needed to produce evidence of proven techniques for reducing salience in future participants.

REFERENCES

- Beaton, A., Tougas, F., Rinfet, N., Huard, N., & Delisle, M., (2007). Strength in numbers? Women and mathematics. *European Journal of Psychology and Education* 22, (3).
- Beilock, S., Rydell, R., & McConnell, A. (2007). Stereotype threat and working memory: mechanisms, alleviation, and spillover. *Journal of Experimental Psychology*, 136.
- Bergeron, D., Block, C., & Echtenkamp, B., (2006). Disabling the able: Stereotype threat and women's work performance. *Human Performance* 19 (2).
- Bonnot, V., & Croizet, J. (2007). Stereotype internalization and women's math performance: the role of interference in working memory. *Journal of Experimental Social Psychology*, 43.
- Bonnot, V., & Croizet, J.C., (2007). Stereotype internalization, math perceptions, and occupational choices of women with counter-stereotypical university majors. *Swiss Journal of Psychology* 66 (3).
- Chang, D., & Demyan, A. (2007) Teacher's stereotypes of Asian black and white students. *School Psychology Quarterly*, 22.
- Chavous, T. (2005). An intergroup contact-theory framework for evaluating racial climate on predominantly white college campuses. *American Journal of Community Psychology*, 36.
- Cohen, G., & Garcia, J., (2005). "I am us": Negative stereotypes as collective threats. *Journal of Personality and Social Psychology* 89, (4).

- Cullen, M., Waters, S., & Sackett, P., (2006). Testing stereotype threat theory predictions for math-identified and non-math identified students by gender. *Human Performance* 99. (4).
- Davis, C., Aronson, J., & Salinas, M. (2006) Shades of threat: racial identity as a moderator of stereotype threat. *Journal of Black Psychology*, 32.
- Davies, P., Spencer, S., & Steele, C., (2005). Clearing the air: safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology* 88.
- Ericsson, K. (1985). Memory skill. *Canadian Journal of Psychology*, 39.
- Gazzaley, A., Sheridan, M., Cooney, J., & Esposito, M. (2007). Age related deficits in component processes of working memory. *Neuropsychology*, 21.
- Goolkasian, P. (2007). Memory for actions rather than things, people, or events: a review of prospective memory. *The Journal of General Psychology*, 134.
- Gresky, D., Eyck, L., Lord, C., & McIntyre, R., (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles* 53, (9/10).
- Huguet, P., & Regner, I., (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology* 99, (3).
- Jamieson, J., & Harkins, S. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, 93.
- Johns, M., Schmader, T., & Martens, A., (2005). Knowing is half the battle: teaching stereotype threat as a means of improving women's math performance. *American Psychological Society* 16.

- Keller, J., (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female student's math performance. *The British Journal of Educational Psychology* 77.
- McGlone, M., Aronson, J., & Kobrynowicz, D. (2006). Stereotype threat and the gender gap in political knowledge. *Psychology of Women Quarterly*, 30.
- McGlone, M., & Aronson, J., (2007). Forewarning and forearming stereotype-threatened students. *Communication Education* 56, (2).
- Muzzatti, B., & Agoli, F., (2007). Gender and mathematics: Attitudes and Stereotype threat susceptibility in Italian children. *Developmental Psychology* 43, (3).
- Osborne, J. (2007). Linking stereotype threat and anxiety. *Educational Psychology*, 27.
- Park, S., Kim, M., & Chun, M. (2007). Concurrent working memory load can facilitate selective attention: evidence for specialized load. *Journal of Experimental Psychology*, 33.
- Risen, J., Gilovich., & Dunning, D. (2007). One-shot illusory correlations and stereotype formation. *Personality and Social Psychology Bulletin*, 33.
- Rosenthal, H., Crisp, R., & Suen, M., (2007). Improving performance expectations in stereotypic domains: Task relevance and the reduction of stereotype threat. *European Journal of Social Psychology* 37.
- Schader, T., Johns, M., & Barquissau, M., (2004). The costs of accepting gender differences: the role of stereotype endorsement in women's experience in the math domain. *Sex Roles* 50.
- Schmader, T., Johns, M., & Forbes, C., (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review* 115, (2).

- Sellers, R., Rowley, S., Chavous, T., Shelton., & Smith, M. (1997). Multidimensional inventory of black identity: A preliminary investigation of reliability and construct validity. *Journal of Personality and Social Psychology, 73*.
- Stangor, C., Carr, C. & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology, 75*.
- Steele, C. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52*.
- Steele, C., (2003). Through the back door to theory. *Psychological Inquiry 14*.
- Steele, C., & Aronson, J., (2004). Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist 59, (1)*.
- Thoman, D., White, P., Yamawaki, N., & Koishi, H., (2008). Variations of gender-math stereotype content affect women's vulnerability to stereotype threat. *Sex Roles 58*.
- Williams, K. (2006). The effects of stereotypes threat on test performance of male and female college students. *College Student Journal, 40*.
- Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology, 33*.

APPENDIX A

20 ITEM SAMPLE OF THE GENERAL RECORD EXAMINATION

ID _____

Solve the following problems 1-12. There is only one correct answer for each of the following questions.

1. If $4x + 12 = 36$, what is the value of $x + 3$?
(A) 3 (B) 6 (C) 9 (D) 12 (E) 18
2. If $7x + 10 = 44$, what is the value of $7x - 10$?
(A) $-6\frac{6}{7}$ (B) $4\frac{6}{7}$ (C) $14\frac{6}{7}$ (D) 24 (E) 34
3. If $4x + 13 = 7 - 2x$, what is the value of x ?
(A) $-10/3$ (B) -3 (C) -1 (D) 1 (E) $10/3$
4. If $x - 4 = 9$, what is the value of $x^2 - 4$?
(A) 21 (B) 77 (C) 81 (D) 165 (E) 169
5. If $ax - b = c - dx$, what is the value of x in terms of a , b , c , and d ?
(A) $\frac{b+c}{a+d}$ (B) $\frac{c-b}{a-d}$ (C) $\frac{b+c-d}{a}$ (D) $\frac{c-b}{a+d}$ (E) $\frac{c-d}{b-a}$
6. If $\frac{1}{3}x + \frac{1}{6}x + \frac{1}{9}x = 33$, what is the value of x ?
(A) 3 (B) 18 (C) 27 (D) 54 (E) 72
7. If $3x - 4 = 11$, what is the value of $(3x - 4)^2$?
(A) 22 (B) 36 (C) 116 (D) 121 (E) 256
8. If $64^{12} = 2^a - 3$, what is the value of a ?
(A) 9 (B) 15 (C) 69 (D) 72 (E) 75
9. If the average (arithmetic mean) of $3a$ and $4b$ is less than 50, and a is twice b , what is the largest possible integer value of a ?
(A) 9 (B) 10 (C) 11 (D) 19 (E) 20
10. If $\frac{1}{a-b} = 5$, then $a =$
(A) $b + 5$ (B) $b - 5$ (C) $b + 1/5$ (D) $b - 1/5$ (E) $\frac{1-5b}{5}$
11. If $x = 3a + 7$ and $y = 9a^2$, what is y in terms of x ?
(A) $(x-7)^2$ (B) $3(x-7)^2$ (C) $\frac{(x-7)^2}{3}$ (D) $\frac{(x+7)^2}{3}$ (E) $(x+7)^2$

12. If $4y - 3x = 5$, what is the smallest integer value of x for which $y > 100$?
 (A) 130 (B) 131 (C) 132 (D) 395 (E) 396

ID _____

Comparison Questions 13-20: Look at the following questions and calculate whether Column A or Column B is more (higher). Circle the answer in the column that has the greater value.

	<u>Column A</u>		<u>Column B</u>
13.	b	$a + b = 13$ $a - b = 13$	13
14.	a	$\frac{2^{a-1}}{2^{n-1}} = 8$	n
15.	x	$4x^2 = 3x$	1
16.	the average (arithmetic mean) of a, b, and c.	$a + b = 1$ $b + c = 2$ $c + a = 3$	1
17.	x	$3x - 4y = 5$ $y = 2x$	y
18.	X	$X/2 - 2 > X/3$	12
19.	the average (arithmetic mean) of r and s.	$3r - 5s = 17$ $2r - 6s = 7$	10
20.	c	$1/c = 1 + 1/d$ c and d are positive	d

Answer Sheet

1. C
2. D
3. C
4. D
5. A
6. D
7. D
8. E
9. D
10. C
11. A
12. C
13. B
14. A
15. B
16. C
17. A
18. A
19. B
20. B

APPENDIX B

BIOGRAPHICAL INFORMATION

ID _____

Age _____

Race: Black___ White___ Asian___ Latino___ Mid-eastern___

Classification: Freshman___ Sophomore___ Junior___ Senior___

Parents Education Level:

High School___ Some College/Associates___ Bachelors___ Graduate/ Professional Degree___

Rate your confidence in mathematics in general:

Below Average___ Average___ Above Average___

Rate the difficulty of the exam you have just taken:

Not Difficult___ Somewhat Difficult___ Difficult___

Have you heard of the stereotype that “males perform better in mathematics than females” prior to today? Yes___ No___

Is Your Major: Mathematics___ Science___ Engineering___ Accounting___ Other___

Was the **FIRST** thing you noticed in the test that there were **no males**? Yes___ No___

Are the majority of your friends Males? Yes___ No___

Do you normally experience test anxiety? Yes___ No___

APPENDIX C

SUBJECT POOL PROPOSAL AD

Subject Pool Proposal Ad

You are invited to participate in a research study. The purpose of the study is to examine differences in an individual's working memory during cognitive tasks. You may participate in the study if you are a UNLV student age 18 or older.

If you volunteer to participate in this study, you will be asked to do the following: complete biographical information and complete a short quantitative assessment. The assessment will consist of items similar to the following: If $3a + 2 = 11$, then $12a + 10$ equals? A) 14 B) 19 C) 144 D) 46. The process may take up to 45 minutes total time.

APPENDIX D

DEBRIEFING

The assessment you have just completed is intended to help identify differences in working memory during cognitive tasks. The primary intent of the test situation was to identify if verbal cues would elevate or decrease the performance of females in any of the three groups. Our goal is to identify if test takers performance is affected by what they are told about the test prior to taking it. You were assigned to one of three groups. Control, Stereotypical or Non-stereotypical. The control group received the test and no additional instructions. The Stereotypical group, were told that males do better than females on the test. The Non-stereotypical group were told that females do better than males. The results of the assessment are completely anonymous. Your personal information is not attached to any of the information in the data. If you have any questions regarding the study you may contact Dr. Rebecca Nathanson, Tarryn McGhie, and the University of Nevada, Las Vegas Office for the Protection of Human Subjects in the event that you have any additional questions or concerns regarding the study you have just participated in.

Dr. Rebecca Nathanson, Associate Professor
Department of Educational Psychology
University of Nevada, Las Vegas
CEB 246
4505 Maryland Parkway, Las Vegas, NV 89154
Phone 702-895-2323
Email: rnathans@unlv.nevada.edu

Tarryn McGhie, Graduate Student
Department of Educational Psychology
CEB 227A
University of Nevada, Las Vegas
4505 Maryland Parkway, Las Vegas, NV 89154
Phone 702-895-3253
Email: mcghiet@unlv.nevada.edu

Office for the Protection of Research Subjects
4505 S. Maryland Parkway
Box 451047
Las Vegas, NV 89154-1047
Phone: 702-895-2794
Fax: 877-895-2794
Email: OPRSHumanSubjects@unlv.edu

APPENDIX E

TABLES 1, 2 AND 3

Table1. Group means and standard deviation

	N	M	sd
Control	31	59.2	22.4
Stereotypical	34	56.6	29.8
Counter-stereotypical	30	49.5	29.1
Total	95	55.2	25.9

Table 2. Group questions attempted and questions correct

	N	M	sd
Control	31		
<i>attempted</i>		12.8	4.7
<i>correct</i>		7.3	3
Stereotypical	34		
<i>attempted</i>		14.1	5.1
<i>correct</i>		7.6	3.8
Counter-stereotypical	30		
<i>attempted</i>		14.4	4.4
<i>correct</i>		7.1	4
Total	95		
<i>attempted</i>		13.7	4.7
<i>correct</i>		7.4	7.3

Table 3. Group means for attempted, correct and the difference between attempted/correct.

	<i>Sum of Squares</i>	df	<i>Mean Square</i>	<i>F</i>	<i>Level of Significance</i>
Correct					
Between Groups	5.137	2	2.568	.193	.825
Within Groups	1226.4	92	13.350		
Total Attempted	1231.537	94			
Total Difference					
Between Groups	44.131	2	22.066	.968	.384
Within Groups	2096.774	92	22.791		
Total	2140.905	94			
Between Groups	46.798	2	23.399	1.450	.240
Within Groups	1484.360	92	16.134		
Total	1531.158	94			

VITA

Graduate College
University of Nevada, Las Vegas

Tarryn McGhie

Degrees:

Bachelor of Arts, Psychology, 2007
North Carolina Agricultural and Technical State University

Master of Science, Educational Psychology, 2010
University of Nevada, Las Vegas

Thesis Title: Verbal Cues: Producing the Same Results in Stereotype Threat Research?:

Thesis Examination Committee:

Chairperson, Rebecca Nathanson, Ph.D.
Committee Member, Gregory Schraw, Ph. D.
Committee Member, Gita Taasobshirazi, Ph. D.
Graduate Faculty Representative, William Cross, Jr., Ph. D.