I CAN(NOT) AVOID DOING BADLY: THE EFFECTS OF PERCEIVED SOURCE OF A SELF-RELEVANT STEREOTYPE ON PERFORMANCE

by

ILAN DAR NIMROD

B. A., The University of Haifa, 2001

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology, Social Psychology Programme

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July, 2004

Library Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

ILAN DAR NIMROD
Name of Author *(please print)*

02/08/2004
Date (dd/mm/yyyy)

Title of Thesis: I CAN(NOT) AVOID DOING BADLY: THE EFFECT OF PERCEIVED SOURCE OF A SELF RELEVANT STEREOTYPE ON PERFORMANCE

Degree: MASTERS OF ARTS Year: 2004

Department of PSYCHOLOGY
The University of British Columbia
Vancouver, BC Canada

Abstract

The theory of stereotype threat states that activating self-relevant stereotypes can lead people to exhibit stereotype-consistent behavior. Stereotype threat most commonly arises under circumstances in which a negative self-relevant stereotype is applicable, the person's membership in the stereotyped group is made salient, and the person believes that their performance on a task will be evaluated.

It seems that a certain element in stereotypes conveys an inescapable expected behavior to members of the stereotyped social group. Putting this assertion to test we manipulated the perceived inevitability of a stereotype-related group difference. Research on Nature vs. nurture causal attributions suggests that people perceive genetic causes to be more inescapable than experiential ones. Using a repeated measures design, causal attributions concerning gender-based differences in mathematical ability were manipulated by presetting either genetic-based or experientially-based explanations for the gender-related math performance differences, while the strength of the alleged differences was held constant. A third condition asserted that there are no gender differences in math. Additional variable tested was the presence of men's influence on women math performance.

Results supported the hypothesis that the perceived cause for gender differences in math ability affects women's mathematical performance. Women who were exposed to a genetic explanation performed significantly worse than those exposed to experiential explanation. Men's presence did not significantly influence women's math performance. The results indicate one way in which genetic essentialism might affect people's behaviour. Several more implications, as well as future directions are discussed.

Table of Contents

Abstract ii

Table of Contents iv

List of Tables vi

List of Figures vii

Acknowledgments viii

Introduction 1
- Stereotypes 3
- Stereotype threat 4
- Suseptability to ST 7
- Implicit theories 9
- Stereotype source 10
- Study's hypothesis 12

Method 14
- Participants 14
- Materials 14
  - Pre-test Measurements 14
  - Test 15
  - Post-test measurements 17
- Design and analysis 18
- Procedure 18

Results 20
- Order effect 20
- Number of questions answered correctly on Math 1 20
- Number of questions answered correctly on Math 2 20
- Accuracy 22
- Additional measures 23
- Mediators 23

Discussion 25
- Limitation and future directions 29
- Conclusions 30

References 31

Figure Captions 42

Appendices 46
Appendix 1A: One version of the math sections 46
Appendix 1B: The second math section 50
Appendix 2A: No gender differences in math abilities manipulation (ND condition essay) 54
Appendix 2B: Males perform better on math tasks due to genetic disposition manipulation (G condition essay) 56
Appendix 2C: Males perform better on math tasks due to experiential circumstances manipulation (E condition essay) 58

List of Tables

Table 1: Scores on math 2 for "source of stereotype" and Setting conditions before and after covariation of Math 1 scores. 41

List of Figures

Figure 1: Math 2 estimated means after controlling for Math 1 scores... 43

Figure 2: The relationship between changes in effect size of "source of stereotype" conditions and accuracy 44

Figure 3: Jayaratne's (2002) suggested model for the relationship between genetic explanations and attitudes towards social groups..................................... 45

Acknowledgments

There are so many people that assisted, guided, and supported me in various stages of this project that I would like to thank. First, I would like to acknowledge all of the participants for the time and energy they invested in the project. I am thankful to Mana Ebrahimi, Jaime Caron, Jennifer Lau, Veronica Tang, and Maria Salampassis for assisting me in countless hours of participant testing and data entry.

I wish to thank my mother for planting the seed of intellectual curiosity, which led me here. I would like to thank my partner Shelby Zuehlsdorff for supporting me relentlessly and providing me with an eager yet critical resonance box for all my half-baked ideas.

I want to thank Ralph Hakstian for his dedication to statistics and his guidance with the appropriate statistical analyses.

I am grateful to the University of British Columbia graduate fellowship and my co-supervisor, Darrin Lehman, for believing enough in my potential to financially support me and my research endeavors. I would also like to thank Darrin Lehman for his continuous guidance and assistance through this educational journey.

I would like to thank my committee member, Ara Norenzayan, for providing (sometimes even unknowingly) guidance and suggested reading that helped form some of the theoretical arguments. Finally, I want thank my co-supervisor Steve Heine for his time and knowledge, guidance, and unparallel contributions to this research and other projects. I am fortunate to have such a group as mentors.

Introduction

In my senior year of high school Mr. Rosenberg, my advanced math teacher, explained to the class that when it comes to math, girls are inherently inferior compared to boys. Bolstering his long, successful career as evidence to back up his assessment, he asserted that women are just built differently, and that their talents lay in other areas; literature, for example. His statement remained uncontested during the rest of the year. Sure enough, though, the grades in the class reinforced his belief, as the boys' grades on average were higher than the girls' grades (a fact, which Mr. Rosenberg did not spare us). Mr. Rosenberg based his argument on his experience as a math teacher, but large scale empirical data shed more objective light on his statement.

In a meta-analysis conducted by Hyde, Fennema, and Lamon (1990) involving over 3 million participants, the findings showed that through elementary and middle school there are virtually no differences on standardized math tests between girls and boys. In the third international mathematics and science survey (TIMSS) (Mullis et al., 2000) 2,668 to 9,981 eighth graders were sampled from 38 countries. The survey results showed significant differences between boys and girls (favoring the boys) in only 4 countries and a very small significant effect for the entire represented international population.

A trend toward better performance of males steadily increases from high school through college (Hyde et al., 1990). In college women have been found to drop out of math, engineering, and physical science comes at a rate two and a half times that of men, which leads to a large disparity between the percentage of women and men holding bachelor and advance degrees in science, math and engineering (Hewitt & Seymour, 1991 as cited in Steele, 1997). In 2003 the number of women in math-related fields is still only a fraction of that of men (Schmader, Johns, & Barquissau, in press).

Studies demonstrated equivalent and even worse findings for the African-American population (Jones, Burton, & Davenport 1984).

A wide variety of factors have been suggested as explanations for the gender and racial discrepancies on academic performance. A bitter discourse on the subject is part of the over arching nature-nurture debate. The genetic camp asserts, as Mr. Rosenberg did, that performance differences are caused by innate differences (e.g., Benbow & Benbow, 1987; Herrnstein & Murray, 1994). The environmental camp, on the other hand, claims that situational conditions present certain groups with disadvantages in developing well needed academic skills (e.g., Eccles & Jacobs, 1987; Jencks, 1998). The debate is not confined to the ivory tower, and the public is deeply influenced by the different views. People seem to form their own opinions on causes of performance differences between social groups (e.g., race, gender) in a way that corresponds to the academic debate (Jayaratne, 2002). Their view of social groups' characteristics is also heavily influenced by an essentialist tendency- perceiving one social group to be fundamentally different from another due to a different "essence" (Rothbart & Taylor, 1992). Nature explanations for group differences resonate with this essentialist tendency.

People's perceptions of differences between social groups seem to match existing stereotypes (Fiske, 1998). The purpose of the present study is to investigate the influence of people's stereotype-related nature-nurture causal attributions on their performance.

Stereotypes

In his seminal book "The nature of Prejudice", Allport (1954/1958/1979, p. 191) defined a stereotype as "an *exaggerated*[1] belief associated with a category". Various studies found associations between stereotypes and different constructs, among them self-esteem (e.g., Greenwald et al., 2002), prejudice (e.g., Lambert, Payne, & Jacoby, 2003), sexual-orientation based aggression (Krahe, 2000), and personality disorders (Reinzi, Forquera, & Hitchcock, 1995). Negative stereotypes have been demonstrated to have detrimental effects on members of the stereotyped groups. Previous research has revealed that stereotypes influence people's actual performance on tasks. Two distinct theories offer theoretical accounts for the process by which stereotypes affect people's behavior (for a comparison discussion of the two see, Wheeler & Petty, 2001).

The *ideomotor mechanism theory* postulates that activating constructs of stereotyped groups by using conscious or subliminal primes can lead to stereotype-consistent behavior (Bargh, Chen, & Burrows, 1996; Dijksterhuis, Aarts, Bargh, & Van Knippenberg, 2000). For example, Bargh et al. (1996) subliminally primed young college students with elderly-related words (e.g., senile, Florida) to activate the elderly construct. Following this procedure they measured the pace by which participants walked down the corridor leaving the laboratory. They found that compared with participants in the control group subliminally primed with neutral words (e.g., table), participants in the experimental condition walked slower. The stereotypes in this case were not targeting groups to which the participants belong. Never-the-less, priming the construct seem to effect the cognition and elicit stereotype-consistent behaviour.

[1] Italics added.

Stereotype threat

The second relevant theory is stereotype threat (ST). It states that activating self-relevant stereotypes leads people to behave in a way that is consistent with the stereotype (Steele, 1997). Numerous studies provide support for ST theory (e.g., Kray, Thompson, & Galinsky, 2001; Levy, 1996; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995), and it has received a great deal of attention in the past decade.

ST most commonly arises under circumstances in which a negative self-relevant stereotype is applicable, the person's membership in a stereotyped group is made salient, and the person believes that their performance on a task will be evaluated. The performed task should be perceived as very difficult (O'Brien & Crandall, 2003, Spencer et al., 1999). Steele (1997) argues that these conditions create an environment in which people tend to act in a way that confirms the stereotype. The main focus of research on ST is performance on academic related tasks. Minority group and gender stereotypes are most commonly addressed in this context. Various studies have demonstrated decreased performance on standardized tests for African-American (Steele & Aronson, 1995), Latin-American (Gonzales et al., 2002), and low SES (Croizet & Claire, 1998) participants. Other studies, focusing on gender-related stereotypes, demonstrated women's decreased performance on math exams (Inzlicht & Ben-Zeev, 2000; Spencer et al., 1999).

Consider for a moment the following scenario. Joe and David are two bright African-American high school students, who wish to apply to prestigious colleges. They study hard in order to excel on the SAT exam and are aware of the importance of their score. Joe is taking the exam in a center in which demographic data (age, race, and marital status) are being collected prior to the exam, where as David is taking the exam in a centre in which the same data is being collected following the exam. Could this minor difference have any effect on their performance? Certain findings suggest that being reminded of his race prior to the exam might hinder Joe from reaching his full ability (Steele & Aronson, 1995, experiment 4).

Although the academic realm is by far the most studied, evidence for a similar process has been found in other areas as well indicating a most prevalent phenomenon. Stone, Lynch, Sjomeling, and Darley (1999) examined the performance of White and Black men in sports. They found that labelling a golf task a measure of "natural athletic ability" led White participants to exhibit decreased performance whereas labelling the task a measure of "sport intelligence" led Black participants to exhibit decreased performance. Another study found that priming elderly people with negative stereotypes of aging, resulted in worsened memory performance (Levy, 1996). Other studies used the paradigm to demonstrate decreased performance of women in a negotiation task (Kray et al., 1998) and decreased performence on an affective word-recognition task for men (Leyens, Desert, Croizet, & Darcis, 2000) This broad array of findings indicates a deeply rooted, domain-general mechanism that hinders performance.

Activation of a stereotype need not be blatant or forceful. Merely reminding a person of his or her membership in a social group that a stereotype is targeting has been shown to be sufficient. A range of subtle ways, which resemble real-world circumstances, has been demonstrated to elicit impaired performance. For example, Steele and Aronson (1995, Experiment 1) presented participants with a test consisting of hard questions from the GRE. The test was presented as a "test that measures verbal ability" (diagnostic condition) or a "test that does not evaluate ability, but provides more understanding about problem solving" (non-diagnostic condition). The results showed that consistent with the stereotype, African Americans in the diagnostic condition performed worse than European-Americans in both conditions. They also performed worse than African-Americans in the non diagnostic condition. The latter performed as well as European-Americans (after controlling for SAT scores). Inzlicht and Ben-Zeev (2000) found that women who took a math exam in the presence of men performed worse than women who took the exam in the company of only other women. This finding echoes previous research that illustrated that minority status can evoke group identity (McGuire, McGuire, & Winton, 1979). Other examples of subtle ways in which stereotypes were activated in previous research are subliminal priming (Levy, 1996), watching television commercials (Davies, Spencer, Quinn, & Gerhardstein, 2002), or completing a demographic questionnaire prior to the test (Steele & Aronson, 1995, Experiment 4).

The implications of a person assimilating the stereotypes are apparent. Stereotyped group members seem to adhere to a negative self-fulfilling prophecy, which undermines the quest to reach their full potential. The ST mechanism perpetuates an undesirable status quo in which the stigmatized live up to the stigma.

Of course, no one argues that ST is the sole reason for the gender- differences found in math performance in surveys (Sackett, Haridson, & Cullen, 2004; Steele & Aronson, 2004). Different scholars give various explanations (i.e., genetic or environmental or a mix) for the performance differences. The common thread for both the genetic and the environmental camps is that they deal with elements that are very hard to change. Making rapid, significant changes in genetic dispositions, poverty, culture, or class is extremely difficult if not impossible, therefore, can offer us little hope of eliminating the discrepancy. Reducing ST, on the other hand, has shown some promising results in reducing the gaps in both laboratory studies (Aronson et al., 1999) and field studies (Aronson, Fried, & Good, 2002).

Susceptibility to ST

Although it is suggested that "...virtually any group can experience it (ST) to a meaningful degree in certain circumstances"(Aronson, 2002, p. 285), a few features have been shown to increase people's vulnerability to ST. These risk factors lead people to show greater underperformance in situations in which self-relevant stereotypes are activated. Features that have been identified include domain identification, group identification, and acceptance of the stereotype.

Domain identification introduces one of the grim ironies of ST. Studies have shown that people who care the most about the task at hand exhibit the greatest decrease in performance in the ST paradigm (Aronson et al. 1999; Aronson & Good, 2001). It has been suggested that individuals who invest the most in a certain domain (e.g., math) will be the ones to feel the greatest pressure to do well, making themselves susceptible to higher levels of anxiety due to the activated negative stereotype. The more important the situation the more pronounced the "choking effect" (Baumeister, 1984).

Group identification is a measure of the level of connection one has to one's group (e.g., race, gender). Research findings seem to suggest that the greater investment individuals have in their group, the more they will be concerned with the stereotypes relating to the group's performance (Schmader, 2002). Rosencrantz (1994, as cited in Aronson, 2002) found that Latin-Americans, who feel grounded in both mainstream culture as well as their origin culture, are less likely to underperform on a standardized test under conditions of ST.

According to Steele and colleagues (Steele, Spencer, & Aronson, 2002), acceptance of the stereotype is not necessary for the ST phenomenon to occur. Recent studies, however, challenge this proposition by indicating that stereotype's acceptance exacerbates impaired performance (Schmader et al., in press). Acceptance of stereotypes by the stereotyped-groups' members, seems to be stronger than was previously supposed. For example, Way (1998) conducted in-depth interviews with minority adolescents. She found that although the adolescents want to reject the negative stereotypes, they seem to be unable to do so confidently. Acceptance of the stereotypes was, also, found to play a role in the relations between gender, stereotypes, and math ability. Studies have shown that parents tend to make gender-based stereotypic attributions to their children's math performance. These attributions were negatively correlated with their daughters' engagement in math-related activities (Eccles et al., 1989). Examining math-related gender stereotypes at a later age, in one sample, 41% of college women stated that there is at least "some truth" to the stereotype of men's higher abilities in math (Blanton, Christie, & Dye, 2002). Using the implicit association test (IAT), women, even ones majoring in math-related fields, endorsed a strong association between math and maleness (Nosek, Banaji, & Greenwald, 2002).

Research on endorsement of stereotypes illustrates heightened vulnerability to ST by people who are more accepting of a self-relevant stereotype (Schmader et al., in press). Schmader et al. (in press) investigated the influence of women's endorsement of the gender-related math stereotype on actual performance on a math related exam. They found that when gender identity was made salient, by telling the participants that the examiner will evaluate their performance as representatives of women in general, women who demonstrated higher stereotype endorsement showed decreased performance compared to both women who did not endorse the stereotype and women who endorsed it but were told that the examiner was interested only in their personal performance. This line of research indicates a greater involvement of endorsement of stereotypes than was previously suggested by Steele et al. (2002).

Stereotypes may convey an inescapable expected behavior to people in the targeted group. People's acknowledgement of their membership in a social group influences their abilities in a stereotype-consistent way. Their performance may be a result of how unavoidable they perceive this stereotyped predicted outcome to be. The present study will assess the role of perceived inevitability in creating stereotype-consistent underperformance. People's implicit theories with regards to stereotypes can direct us in this line of inquiry.

Implicit theories

Dweck and colleagues (Dweck, Chiu, & Hong, 1995; Hong, Chiu, Dweck, & Sacks, 1997; Mueller & Dweck, 1998) examined implicit theories that people hold about the malleability of intelligence and morality. They categorized as "entity theorists" people who perceive traits as fixed and as "incremental theorists" people who perceive traits as malleable. Regardless of the absolute boundaries of people's intelligence (Sternberg, 1985) and abilities, following the entity vs. incremental classification, one can assume that entity theorists are particularly inclined to seek out trait information for its predictive value in evaluating others. Incremental theorists, on the other hand, should believe that traits are more dynamic, and people's behavior may change over time and across situations. Information about traits, therefore, does not provide as much predictive power for incremental theorists. Empirical evidence provides support for these assumptions (Chiu, Hong, & Dweck, 1997; Erdley & Dweck, 1993).

Stereotypes often include assumptions about traits, which can be used as a source of information. Stereotypes often play a leading role in evaluative judgments (Kunda & Sherman-Williams, 1993). The relation between implicit theories and the processing of stereotype-relevant information has been the focus of many recent studies (e.g., Plaks, Stroessner, Dweck, & Sherman, 2001; Sousa & Leyens, 1987; Wittenbrink, Hilton, & Gist, 1998; Yzerbyt, Leyens, & Corneille, 1998). Plaks et al. (2001) found that entity theorists directed more attention towards stereotype-confirming information whereas incremental theorists showed either no preference or preference for stereotype-disconfirming information. Furthermore, entity theorists have been found to endorse stereotypes more than incremental theorists (Levy, Stroessner, & Dweck, 1998, Experiment 1) and even experimentally induced entity thinking led participants to endorse stereotypes more strongly (S. Levy et al., 1998, experiment 4). Thus, stereotypes seem to convey a different message to entity theorists than to incremental theorists.

Stereotype source[2]

Levy et al. (1998, Experiment 2) asked participants to provide an explanation for the existence of several African-American stereotypes. The participants had to indicate how much they agreed with different statements that provided inherent or situational explanations of stereotypical views of African-Americans. Entity theorists judged the stereotypes to originate from innate/inherent characteristics in which African-Americans differ from European-Americans. On the other hand, incremental theorists tended to endorse situational factors, faced by African-Americans to explain the stereotype origin. These findings strengthen the view that people perceive inherent characteristics as harder (if not impossible) to change, whereas situational factors are judged to be more pliable and less restrictive.

[2] For the purpose of this study, I do not draw a conceptual distinction between source, origin, or cause of a stereotype. I use them interchangeably to refer to underlying explanation of differences between social groups that are specified by a stereotype.

One can find this notion endorsed by many people, including researchers, as self-explanatory and self-evident (Nelkin & Lindee, 1995). For example, Steele (1997, p. 624) describes a study in which he and Stoutemeyer varied the strength of stereotype threat. The manipulation they used attributed math performance differences between women and men to a “small but stable innate ability…or to social causes such as sex-roles prescriptions and discrimination”. The findings of this study were that “women under stronger stereotype threat disidentified with math and math related careers more than women under weaker stereotype threat”. Steele found no reason to explain or even identify which of the conditions (innate or social) serves as the “stronger stereotype threat” undoubtedly assuming the readers’ recognition of the innate condition as the stronger manipulation.

There is further empirical evidence suggesting that people judge genetic pre-disposition as less changeable than disposition acquired through experience. For example, Monterosso, Royzman, and Gleitman (2004) found that undesirable behaviors that were explained with reference to people’s somatic (e.g., genetic, head injury) pre-dispositions rather than experiential (e.g., trauma, abuse) causes were seen as less voluntary, seen as less blameworthy, attached greater sympathy when the explained behavior led to adverse consequences, and were assigned less punishment. In addition, participants felt that they were more likely to have behaved similarly to the people in the vignettes if they had shared the somatic pre-disposition but not if they had shared the same instigating experiences. The tendency to undermine the importance of situational factors compared with dispositional ones in evaluating undesired behaviors is not restricted to laypeople. Judges have been found to disregard significant situational factors, and tend to take no notice of external constraints even when imposed by themselves (Gilbert & Jones, 1986).

Study's hypotheses

Allport (1954) identified stereotypes as "exaggerated" belief. Following Allport's conceptualization, exaggeration may be embedded in people's implicit beliefs that stereotypes targeting their own group come from a genetic pre-disposition, a hard-wired, inescapable, built-in difference between their group and the non-stereotyped one. This believed unavoidable difference can be the causal force behind their reduced performance on stereotyped-related tasks. Members of stereotyped groups are more inclined to believe their abilities are derived from the genetic pre-disposition common to their group, pre-disposition that can be seen as a marker of their group membership.

To examine this hypothesis I manipulated the perceived source of the stereotype, giving either a genetic or experiential/environmental explanation for "observed" (stereotyped) gender differences in mathematical ability. A third condition asserted that there are no gender differences in math abilities. Asserting no gender differences was demonstrated in the past to eliminate ST-related underperformance (Spencer et al., 1999). The main hypothesis of the study was that women who are led to believe that the stereotype "women do worse in math than men" originates from a genetic disposition will show worse performance than both women who believe that the stereotype originates from an experiential cause (teachers' expectations) or women who are led to believe that there are no math-related gender differences.

The study is also aimed at extending Inzlicht and Ben-Zeev (2000) findings that presence of men leads women to show decreased performance on math tasks to more natural settings. Inzlicht and Ben-Zeev found this phenomenon to occur in small groups performing the task in a lab. The current study tested the influence of the presence of men on women's math performance in larger groups composed of class mates and in a classroom setting. This manipulation increases the ecological validity of the findings.

Additional measures will examine cognitive and emotional variables. Anxiety was previously demonstrated to mediate ST findings (e.g., Osborne, 2001) though it failed to produce conclusive findings in others (e.g., Aronson et al., 1999). Self-handicapping also mediated ST in some studies (Keller, 2002; Stone, 2002) but not in others (Croizet & Claire, 1998). Steele and Aronson (1995) found that African-Americans under ST avoided conforming to stereotypic images of their social group by distancing themselves from a variety of activities, music, sports, and personality dimensions stereotypically associated with their race (stereotype avoidance). I hypothesize that anxiety and self-handicapping will be higher in the genetic condition due to sense of looming inescapable undesired outcome. These construct may potentially mediate the ST effect. I also hypothesized that participants in the genetic condition will show greater stereotype avoidance behavior due to the more immanent threat to their perceived abilities.

Method

Participants

Eighty-nine female participants and twenty males were recruited from the UBC student population. Psychology students received course credit for their participation whereas students in other departments participated in exchange for two movie passes. I was interested in the performance of women in math; therefore the analyses examine only women's data. Fourteen participants were dropped from the central analysis, due to performance on the first math test (prior to the manipulation) that was below chance. This performance indicates an insufficient level of math for the purpose of the study. The participants were run in either small groups (2-6) of females only to reduce "minority status" effect (Inzlicht & Ben-Zeev, 2000) or larger groups (8-15) containing both males and females to increase ecological validity by providing conditions that are mostly common in regular education settings. Participants were 17-39 years old (mean=20.5, SD=3.0).

Materials

Pre-test measurements

A one item likert scale tapped motivation (how motivated are you to do well on the exam?). The scale ranged from 1 (not motivated at all) to 7 (extremely motivated).

A one item likert scale assessed expectancies (how well do you think you will do on the math section(s)?). The scale ranged from 1 (very badly) to 7 (very good).

Three items were administered to measure identification with math (e.g., how much do you care about being good in math?). The scale ranged from 1 (not at all) to 10 (very much). There were no reversed items and the cumulative score on the scale was taken as the measure.

Test

The test was designed to resemble a GRE template. The test was constructed of two sections of math, each containing eighteen difficult questions taken from a pool constructed by the author as study materials for the Israeli psychometric test (equivalent to the American SAT) when he worked as a psychometric test preparation course instructor (see appendix 1A, 1B). These section ordering was counterbalanced.

A verbal section including the manipulation was inserted in between the math sections. The verbal section contained analogies, sentence completion, and reading comprehensions that were shown as extracts from The New York Times on the web to increase the reliability of the arguments embedded in them. Three different reading comprehensions served as the "source of stereotype" manipulation. The bogus New York Times articles claimed that researchers were able to find one of the following results:

1. No gender differences in math abilities (ND condition, appendix 2A). A representative paragraph from this essay read as follows:

> Using the largest sample (over 50 million people!) ever recorded the researchers compared grades in mathematics and physics in 113 countries around the world. The grades were taken from national comparison exams in elementary schools and high schools. The results showed that males and females were performing just as well on the math sections. The same results were found among college students in 26 countries...

2. Males perform better on math tasks due to genetic disposition (G condition, Appendix 2B). A representative paragraph from this essay read as follows:

> The new research is the largest published study of polygenetic effects to test the interaction between different genes and higher cognitive functions. One of the main findings demonstrates an interaction of 2 genes located on the Y chromosome (which is found only in males) with genes

on chromosome 5 and chromosome 7. This interaction produces hormonal changes guided by the hypothalamus. The onset of the hormonal release is guided by activation of the Brotically area in the frontal lobe. This area is activated when processing mathematical oriented problems. F-MRI scans show these hormonal changes create an increase in the amount of ATP (the body's currency of energy) molecules directed to the hippocampus when a person is engaged in higher mathematical reasoning tasks. The increased energy to this area of the brain, considered the "working memory organ", enables the person to retain more accessible short term memory information while concentrating, a critical element in mathematical reasoning capabilities. This genetic difference seem to explain the findings that boys show superior performance by having on average a grade 5 percentile points higher than girls...

3. Males perform better on math tasks due to experiential circumstances (E condition, Appendix 2C). A representative paragraph from this essay read as follows:

The new research is the largest published study of differences among males and females in mathematical reasoning. The research was conducted over 8 years in which the participants were followed and their performance closely observed. Unlike previous research in the field, the present study followed both a genetic research design (to look for inherit differences) and a cognitive research design, which used teachers expectations and behaviour. In the genetic paradigm, using top of the line instruments (F-MRI, DNA analyzers, and messenger RNA blockers) the researchers failed to find any gender differences on mathematical tasks... In the experimental condition the researchers visited schools as educational psychologists and gave the students a bogus mathematical test at the beginning of the year then providing the teachers with fake reports that illustrated that the girls in the class were better in mathematics. Observing the teachers through a video camera in the class it became apparent the teachers were paying more attention to the girls, were more praising towards them and were more dismissive of the boys. In the control condition, where no manipulations of teachers' expectations had taken place, the opposite pattern was observed... The findings showed that the girls in the experimental condition were superior to the boys if the teachers' expectations were manipulated ... In the control conditions, boys showed superior performance by having on average a grade 5 percentile points higher than the girls through out the 8 years of the experiment, providing more support to the general stereotype...

All the manipulations were given in the early part of the verbal section to insure participant's exposure to the manipulation. All the participants answered the questions about the essays that contained the manipulation. An additional unrelated reading comprehension essay was embedded in the verbal section to decrease suspicion.

Post-test measurements

One item likert scales ranging from 1 to 7 were administered to assess perceived difficulties (Did you have any difficulties answering the math questions? 1-not at all, 7-very much), effort (How much of an effort did you make to answer the math questions? 1- I didn't put any effort, 7- I put a lot of effort), and subjective performance evaluation (How well do you think you performed on the math section(s) compared to other participants? 1- much worse, 7- much better). To reduce suspicion for each question that addressed math there was an equivalent question addressing verbal reasoning.

To assess stereotype avoidance participants were asked to indicate on a 1-7 scale how much they like different activities and sports (1- not at all, 7- very much), and how much certain personality dimensions describe them (1- does not describe me at all, 7- describes me very well). Eight activities were selected, four of them stereotypically associated with women (e.g., shopping) and four unrelated to gender (e.g., travelling). Nine sports were mentioned and as with activities, four were stereotypically feminine (e.g., figure skating) and five were either unrelated to gender or masculine (e.g., basketball). Nine traits were given in the same format, five were stereotypically feminine (e.g., emotional) and four unrelated to gender (e.g., humorous).

Self-handicapping was assessed on a two item scale (Steele & Aronson, 1995) ranging from 1 to 7 (e.g., How much stress have you been under lately) . State anxiety was measured on the 20 item State-Trait Anxiety Inventory (STAI) (Spielberger, 1970) rated on a scale ranging from 1 to 4 (e.g., I am tense). This questionnaire was followed by surveying demographic data.

Design and Analysis

Women in the study participated in all experimental conditions whereas men served as controls and participated in one condition only. A 3 (G condition, E condition, ND condition) x 2 (single sex group, mixed sex group) between-subjects repeated measures design was used to assess ST effects. Females were randomly assigned to one of the three "source of stereotype" conditions. To increase ecological validity the participants in the mixed group condition were recruited in classes and were administered the materials in a classroom setting with their peers. The men in the study were part of the mixed-sex groups and M (male) condition, which did not include the gender related reading comprehension. For each participant, the scores on the first math section (pre-manipulation) were taken as their math ability's base rate. They served as a covariate in the analysis of the dependent variable- scores on the second math section (post-manipulation). The scores were examined in GLM analysis of covariance and were compared between conditions.

Procedure

Participants were recruited for an experiment called "Improving grad schools' acceptance criteria". They were told that the researchers wee attempting to improve ecological validity of the current Graduate Record Exam (GRE), which is used as one of the most important criteria for acceptance to graduate school. They were also told that they would be asked to take a GRE style exam and complete a few questionnaires.

Upon arriving at the exam rooms participants were greeted by one of four female experimenters. After signing a consent form, they were asked to complete the pre-test questionnaire. Following the pre-test they were administered a test package that contained sample questions of math and verbal reasoning, which the experimenter used to demonstrate what they would be asked to do on the exam. Following the sample questions they were given 15 minutes to complete each section with no intervals between them. All the participants started with a math section (Math 1) followed by a verbal section and then with a second math section (Math 2). To further reduce suspicion they were told that the order and the number of each sections were randomly distributed (i.e., a person might get two verbal sections and one math section or vice versa).

Following the exam participants were asked to complete the post-test and demographic questionnaires. After completing the post-test questionnaires, participants were thoroughly debriefed about the true purpose of the study and were thanked for their participation.

Results

Order effect. To rule out an order effect, an independent samples t-test was performed on the dependent variable with the order of the two sets of math tests as an independent variable. Results showed no sign of order effect ($T_{1,87}$= 1.35, p>.1), suggesting an equivalent level of difficulty of the two math sets used.

Number of questions answered correctly on Math 1. Different conditions might have differed on their initial math ability. An analysis of variance (ANOVA) on the number of questions answered correctly on the first math exam, however, did not reveal significant differences between conditions ($F_{2,72}$=1.30, p=.279).

Number of questions answered correctly on Math 2. To reduce noise due to differences in initial math abilities, subsequent analyses controlled for differences in Math 1 scores. An analysis of covariance (ANCOVA) on the number of questions correctly answered in the second math exam (see Table 1) revealed significant differences for "source of stereotype" ($F_{2,68}$=4.53, p=.014), no significant differences for setting ($F_{1,68}$=2.48, p=.12), and no significant differences for the "source of stereotype" *setting interaction ($F_{2,68}$=.71, p=.495) [3].

♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦

Insert Table 1 about here

♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦♦

[3] Analysis of covariance (ANCOVA) on the number of questions correctly answered in the second math exam by the entire sample (including women who perform at chance level or worse) revealed significant differences for "source of stereotype" ($F_{2,82}$=3.1, p=.05), no significant differences for setting ($F_{1,82}$=2.31, p=.13), and no significant differences for the "source of stereotype" *setting interaction ($F_{2,82}$=.47, p=.63).

Presence of men did not have significant effect on women's initial math performance (Math 1) ($F_{1,73}$=1.17, p=.24). On math 2 the women only group (M=4.85, SE=.30)* performed slightly better (no significance) than the women who were in the presence of men (M=4.22, SE=.26)*, yet the study failed to replicate previous research (Inzlicht & Ben-Zeev, 2000) at the conventional significance level of .05.

To further investigate the differences between "source of stereotype" conditions, a generalization of the Tukey honestly significant difference procedure due to Bryant and Paulson (1976) revealed the following results for the pairwise multiple comparisons (see Figure 1); participants who were led to believe that men outperform women in math due to a genetic difference performed significantly worse than participants who were led to believe that there are no gender differences in math ability ($qBP_{1,3,71}$=-3.48, p<.05), and significantly worse than participants who were led to believe that men outperform women due to an experiential difference ($qBP_{1,3,71}$=-3.66, p<.05). No significant differences were observed between the E condition and the ND condition ($qBP_{1,3,71}$=-.06, p>.50).

♦♦

Insert Figure 1 about here

♦♦

---

* Estimated means after controlling for Math 1 scores

Accuracy. The ratio between number of questions answered correctly ($Q1_{score}$) and the number of questions attempted ($Q1_{attempted}$) on Math 1 suggests that participants engaged in a large number of guesses (e.g., 14 out of 89 participants (13%) performed at chance level or worse and only 20 out of 89 participants (22%) correctly answered half or more of the questions they attempted). Guesses introduce noise to the experiment that might reduce the effect size for the influence of "source of stereotype" manipulation on Math 2 performance. To address this possibility I calculated the effect sizes of the "source of stereotype" variable for different levels of the ratio $Q1_{score}$/ $Q1_{attempted}$. I then correlated these ratios with their matching effect sizes (see Figure 2) and found positive correlation between them ($r_{Q1score/\ Q1attempted,\ \eta^2}$ =.70, p<.001). this illustration suggests that participants' guessing (smaller $Q1_{score}$/ $Q1_{attempted}$) dampened the effect size of "source of stereotype" on Math 2 performance in the current study.

♦♦

Insert Figure 2 about here

♦♦

One can argue that the increased effect size is due to higher level of identification with math for the people who are more accurate, making them more susceptible to ST. This notion suggests an alternative explanation, other than guessing, for the correlation between $Q1_{score}$/ $Q1_{attempted}$ and the effect size. To examine this explanation I correlated identification with math (IDM) scores with $Q1_{score}$/ $Q1_{attempted}$. The data do not support this alternative explanation ($r_{Q1score/\ Q1attempted,\ IDM}$=.15, p>.10).

Additional measures. Certain initial differences between the "source of stereotype" conditions could have an effect on math 2 scores. Participants' motivation (Kray et al., 2001) and expectancies (Shih et al., 1999) were both suggested as potential mediators for ST findings. Empirical examination of these variables suggests that the effect did not emerge due to initial differences in motivation ($F_{2,72}$=2.3, p=.11, The E condition participants showed the lowest motivation), expectancies ($F_{2,72}$=.63, p=.53).

Steele and Aronson (1995) found that participants who faced ST manipulation showed stereotype avoidance. An ANOVA on stereotype avoidance scores in the different "source of stereotype" conditions revealed marginally significant differences ($F_{2,72}$=2.93, p=.06). Contrary to my hypothesis, Tukey honestly significant difference analyses indicate that the differences originate from higher avoidance in the E condition (M=53.30) than the ND condition (M=58.85) at a p=.05 level[4]. The G condition did not significantly differ from the other two (M=56.07).

No significant differences between "source of stereotype" conditions were found in perceived difficulties on math ($F_{2,71}$=1.75, p>.10) and on math performance evaluation ($F_{2,71}$=.416, p>.6) after controlling for Math 1 scores.

Mediators. Previous research has found evidence that certain variables mediate ST effects experienced by women taking a math exam. The current study incorporated the following three variables that had been identified as potential mediators; anxiety (Spencer et al., 1999) and self-handicapping (Keller, 2002). Effort (Aronson et al., 1999) had been theoretically suggested as a mediator, and was included as well. However, regression analyses revealed no significant mediation roles for any of these variables (all P's>.10). Thus, although women in the G condition showed a ST effect, their

[4] Lower scores indicate lower identification with stereotypically women's characteristics.

impaired performance was not paralleled neither by significant changes in their self-reported anxiety, or self-handicapping, or their effort.

Discussion

I identified an inclination in past research to see attribution due to innate predispositions as stronger than attribution due to experiential ones as a truism (Steele, 1997; Levy et al., 1998). Furthermore, empirical findings support this notion (Monterosso et al., 2004). I harnessed this mechanism to manipulate people's perceptions of inescapability from a self-relevant stereotype-related outcome. We found that women who perceive the stereotype "men are better at math than women" to originate from group differences in genetic make-up exhibit more stereotype-consistent performance on a math task than women who perceive the stereotype to originate from collective gender-differences in related experiences. Perception of stereotype source, therefore, has an effect on stereotype-related task performance, suggesting that stereotyped group members experience stereotype-related performance prediction as inescapable.

Steele and Aronson (1995) found that African-Americans exhibit stereotype avoidance behavior under ST conditions. The present study found that contrary to the original hypothesis participants in the E condition rather than the G condition demonstrated stereotype avoidance. One possible explanation for this finding is that participants in the E condition are actively trying to reduce their similarity to stereotypic exemplars of their gender, and by that they are able to overcome the threat to their performance. Having an experience-based explanation to gender-related math discrepancies externalizes the reason for the discrepancies and thus may implicitly suggest that by distancing herself from women's traditional experiences one can avoid the associated math performance decline. Participants in the G condition, on the other hand, are bound to their group in a genetic-essentialist way, and therefore avoidance of the group characteristics (even stereotypical ones) is futile.

There were no significant performance differences between women who were led to believe that women are inferior in math due to experiential reasons and women who were led to believe that there are no gender differences in math. This evidence suggests that the reason that women show underperfomance following the traditional ST manipulation (exposing participants to a gender prime) may be due to a tendency to naturally perceive their math ability to be a hard-wired, biologically determined ability associated with their social group. It also suggests that this natural inclination towards biological essentialism can be overridden by an explicit environmental explanation. One can offer an alternative explanation for this finding. It can be argued that the experiment lacked sufficient power to uncover a more subtle difference between these two conditions. This alternative explanation is not supported by the data as the difference between the E and ND condition after controlling for initial math ability was minuscule. A replication of this study is needed to allow us to argue more confidently that perceiving the source of a stereotype to be experiential eliminates stereotype threat effect.

Inzlicht and Ben-Zeev (2000) found that women demonstrate reduced performance on a math task in the presence of men. Their study provided women with a highly artificial setting. They found that when put in groups of three in a lab, women showed underperformance when at least one of the people in the group was a man. We failed to replicate this result in a more ecologically valid setting- a classroom with both men and women. The current study's findings suggest that the effect size of ST due to minority status might be lower in more natural environments than the one reported by Inzlicht and Ben-Zeev (Cohen d=.73 in Experiment 1).

How does a manipulation of stereotype source reduce ST effects? Internal beliefs have been shown to predict vulnerability to ST (Inzlicht & Ben-Zeev, 2003;

Schmader, 2002; Schmader et al., in press). The source of the stereotype and the stereotype's effect on the individual constitute a part of a person's internal belief system. Providing an experiential explanation for a stereotyped group performance decrement allows people to assess whether they have encountered similar situations and/or dismiss the effect of the situation on them as unlikely or inapplicable. Lay perceptions of genetic effects are much more difficult to elude (Jayaratne, 2002; Monterosso et al., 2004). A significant percentage of people identify genes as the source of abilities' potentials (Laine et al., in press) and as underlined person's essence (Nelkin & Lindee, 1995). More relevant to this study, a representative survey of Americans found that 57% of the people believe that individual math ability is influenced or determined by genes, and 40% believe that gender differences in math ability are influenced or determined by genes (Jayaratne, 2002). Social desirability concerns (avoiding an undesired sexist label) suggest viewing these numbers as conservative estimate. Taken together with evidence that shows acceptance of the stereotype (or even failure to reject it completely) to be an important risk factor (Schmader et al., in press), these findings offer a deeper understanding of the importance of perceived inescapability in the underlying process that govern ST. The present study suggests that women consider their gender in an essentialist way in the ST paradigm. Their womanhood is part of their essence and therefore this influences their abilities by reducing malleability. Future research should target the role of people's implicit theories (entity vs. incremental) in their susceptibility to ST.

The present methodology also raises an interesting and underinvestigated area. Lay people's perceptions of genes have been shown to be full of misconceptions (Laine et al., in press). These perceptions, however, are marked by essentialism (i.e., genetic essentialism). This essentialism was not created in a vacuum; several mechanisms may have been involved in its formation.

Philosophy of biology theorists assert that there is a basic human tendency towards biological essentialism (also known as folk biology), which can be defined as seeing a living thing as having a certain unchangeable underlying characteristics, whereas artifacts do not. This theory has received a wide range of support from research in developmental psychology (e.g., Gelman, 2003; Keil, 1989) and anthropology (Ahn et al., 2001; Atran, 1987, Atran et al., 2001). Essentialist thinking dominates both physical characteristics (e.g., Gelman, 1988; Keil, 1989) and psychological traits (e.g., Heyman & Gelman, 2000; Sousa, Atran, & Medin, 2002). Medin and Ortony (1989) claimed that people's concepts are often derived from essentialist implicit theories. In support of this claim, several studies indicate that social groups are perceived to follow the essentialist route (folk sociology) in a similar fashion to living organisms rather than man made artifacts, although epistemologically they should be categorized as the latter (Rothbart & Taylor, 1992).

Biological essentialist thinking seems to precede the ability to reason about it (Heyman & Gelman, 2000; Johnson & Solomon, 1997). The concept of genetics may provide the scientific justification for embedded biological essentialism, as genes are seen in part as human essence, although a rigorous scientific approach does not support many of the general public's genetic assertions. Even children adopt a genetic rationalization to justify nature-based essentialist views of psychological traits. For example, explaining a nature-based response to a question about the smartness of a child that was born to "not so smart" parents but was raised by "smart" parents, some fourth and fifth graders provided the following answer "*It will have trouble. It's in his genes*" (Heyman & Gelman, 2000, p. 672). Thus, genetic essentialism might be an extension (or a mature form) of the basic human tendency towards biological essentialism.

Jayaratne (2002) suggested a model (Figure 3) in which genetic explanations for individual differences lead to genetic explanations for perceived gender, class or race differences, which in turn lead to attitudes toward women, the poor or blacks. The results of the current study suggest that the implications of genetic explanations for perceived gender differences (which might generalize to class and race as well) might lead not just to attitudes but also may affect performance on stereotype-related tasks.

♦♦

Insert Figure 3 about here

♦♦

Limitations and future directions

This study introduced a novel manipulation designed to create a stereotype-related underperformance. As with all single study designs, the findings should be replicated and the stereotype source manipulation that was used to produce ST should be compared to the traditional ST manipulation that activates the stereotype subtly, by priming people with their membership in a stereotyped group.

Another limitation of this study was brought about by the level of difficulty of the math sections, which served as the covariate and the dependent variable. In order to create a ST underperformance, the task (DV) should be very difficult. The math sections that I used, though, proved to be too difficult resulting in increased numbers of guesses, dropping participants from the central analyses, and potential floor effect, which may have lowered the effect size for the variables that were studied. Future studies should use easier math questions to avoid these undesirable outcomes.

This study also raises intriguing follow-up questions. For example, how closely related are genetic-based reasoning and stereotype endorsement? Jayaratne (2002) demonstrates that people who hold genetic explanations for perceived gender, class or race differences endorse stereotypical attitudes more. Taken together with the current study, one wonders if people who are primed with the concept of genetics (e.g., genes determine one's love for four-legged animals) might show greater endorsement of unrelated stereotypes (e.g., Jewish people are greedy) due to an increase in essentialist thoughts. Findings of this sort may call for a fundamental change in the way that findings in the genetic field should be reported to the public.

The interaction between stereotypes and genetic essentialism provides a different avenue to explore such as whether priming stereotypes increase genetic essentialism.

Conclusion

Research is just scratching the surface of laypeople's perceptions of genes. The current study suggests one way in which genetic essentialism might affect people's behavior. This line of research is increasing in importance as perception of genetics is progressively more involved in our daily life from increasing media coverage, impact on judicial verdicts, reproductive decisions, and health (for review of these genetic-related implications and more, see Nelkin & Lindee, 1995). This study is only the first step directed towards understanding how perception of genetic causation can shape our behavior.

References

Ahn, W., Kalish, C. W., Gelman, S. A., Medin, D. L., Luhmann, C., & Atran, S. (2001). Why essences are essential in the psychology of concepts. *Cognition*, 82, 59-69.

Allport, G. W., (1954/1958/1979). *The nature of prejudice*. Reading, MA, USA: Addison-Wesley.

Aronson, J. (2002). Stereotype threat: Contending and coping with unnerving expectations. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 279-301). San Diego, CA: Academic Press.

Aronson, J., Fried, C., & 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.

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

Atran, S. (1987). Ordinary constraints on the semantics of living kinds: A commonsense alternative to recent treatments of natural-object terms. *Mind and Language*, 2, 27-63.

Atran, S., Medin, D. L., Lynch, E., Vapnarksy, V., Ek, E., & Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatek Maya in cross-cultural perspective. *Journal of Cognition and Culture*, 1, 3-43.

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

Baumeister, R.F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skilful performance. *Journal of Personality and Social Psychology*, 46, 610-620.

Benbow, C. P., & Benbow, R. M. (1987). Extreme mathematical talent: A hormonally induced ability? *In* D. Ottoson (Ed.), *Duality and unity in the brain* (pp. 147-157). London: Macmillan.

Blanton, H., Christie, C., & Dye, M. (2002). Social identity versus reference frame comparisons: The moderating role of stereotype endorsement. *Journal of Experimental Social Psychology*, 38 , 253-267.

Bryant, J. L. & Paulson, A. S. (1976). An extension of Tukey's method of multiple comparisons to experimental designs with random concomitant variables. *Biometrika*, 63, 631-638.

Burger, J. M. (1991). Changes in attributions over time: The ephemeral fundamental attribution error. *Social Cognition*, 9, 182-193.

Chiu, C., Hong, Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, 73, 19-30.

Church, A. T., Ortiz, F. A., Katigbak, M. S., Avdeyeva T. V., Emerson A. M., Flores J. J. V., & Reyes J. I. (2003). Measuring individual and cultural differences in implicit trait theories. *Journal of Personality and Social Psychology*, 85, 332-347.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96, 608-630.

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

Croizet, J-C., Desert, M., Dutrevis, M., & Leyens, J-P. (2001). Stereotype threat, social class, gender, and academic under-achievement: When our reputation catches up to us and takes over. *Social Psychology of Education*, 4, 295-310.

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

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.

Dijksterhuis, A., Aarts, H., Bargh, J. A., & Van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, 36, 531-544.

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

Dweck, C. S. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 38-61). San Diego, CA: Academic Press.

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, 6, 267-285.

Eccles, J. S., & Jacobs, J. E. (1987). Social forces shape math attitudes and performance. In M. R. Walsh (Ed.), *The psychology of women: Ongoing debates* (pp. 341-354). New Haven CT: Yale University Press.

Eccles, J., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, 21, 215-225.

Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: Relations and changes at early adolescence. *Journal of Personality*, 57, 283-310.

Erdley, C. A., & Dweck, C. S. (1993). Children's implicit personality theories as predictors of their social judgments. *Child Development*, 64, 863-878.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.

Gelman, S. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, 20, 65-95.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought.* New York: Oxford University Press.

Gilbert, D. T., & Jones, E. E. (1986). Perceiver-induced constraint: Interpretations of self-generated reality. *Journal of Personality and Social Psychology*, 50, 269-280.

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.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.

Herrnstein, R. J., & Murray, C. A. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of novel predicates to make inductive inferences about people. *Cognitive Development*, 15, 263-280.

Hong, Y. Y., Chiu, C. Y., Dweck, C. S., & Sacks, R. (1997). Implicit theories and evaluative processes in person cognition. *Journal of Experimental Social Psychology*, 33, 296-323.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta analysis. *Psychological Bulletin*, 107, 139-155.

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.

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology*, 95, 796-805.

Jayaratne, T. E. (2002, June). White and Black Americans' genetic explanations for perceived gender, class and race differences: The psychology of genetic beliefs. Invited lecture at the 2002 Human Genome Lecture Series, National Human Genome Research Institute, NIH, Bethesda, MD.

Jencks, C. (1998). Racial bias in testing. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 457-479). Washington, DC: Brookings Institution.

Johnson, S. C., & Solomon, G. E. A. (1997). Why dogs have puppies and cats have kittens: The role of birth in young children's understanding of biological origins. *Child Development*, 68, 404-419.

Jones, L. V., Burton, N. W., & Davenport, E. C. (1984). Monitoring the mathematics achievement of Black students. *Journal for Research in Mathematics Education*, 15, 154-164.

Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Keller, J. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, 47, 193-198.

Krahe, B. (2000). Sexual scripts and heterosexual aggression. In T. Eckes & H. M. Trautner (Eds.). *The developmental social psychology of gender* (pp. 273-292). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

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

Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin*, 19, 90–99.

Laine, A. D., Jayaratne, T. E., Sheldon, J. P., Kardia, S. L. R., Anderson, E. S., Feldbaum, M., & Petty E. M. (in press). Exploring the public understanding of basic genetic concepts. *Journal of Genetic Counseling.*

Lambert, A. J., Payne, B. K., & Jacoby, L. L. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, 84, 277-295.

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

Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74, 1421-1436.

Leyens, J.-P., Desert, M., Croizet, J.-C., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin*, 26, 1189-1199.

McGuire, W. J., McGuire, C. V., & Winton, W. (1979). Effects of household sex composition on the salience of one's gender in the spontaneous self-concept. *Journal of Experimental Social Psychology*, 15, 77-90.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-196). New York: Cambridge University Press.

Miller, A. G., & Lawson, T. (1989). The effect of an informational option on the fundamental attribution error. *Personality and Social Psychology Bulletin*, 15, 194-204.

Monterosso, J. R., Royzman, E., & Gleitman, H. (2004). Explaining away volition: The effects of scientific explanation on perceived responsibility. *Unpublished manuscript*, University of Pennsylvania.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75, 33-52.

Nelkin, D., & Lindee, M. S. (1995). *The DNA mystique: The gene as a cultural icon*. N.Y: W. H. Freeman & Co.

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

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

Osborne, J. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, 26, 291-310.

Oswald, D. L., & Harvey, R. D. (2001). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social*, 19, 338-356.

Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology*, 80, 876-893.

Rienzi, B. M., Forquera, J., & Hitchcock, D. L. (1995). Gender stereotypes for proposed DSM-IV negativistic, depressive, narcissistic, and dependent personality disorders. *Journal of Personality Disorders*, 9, 49-55.

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & Fiedler, K. (EDs), *Language, interaction and social cognition* (pp. 11-36). Thousand Oaks, CA: Sage Publications, Inc.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American-White differences on cognitive tests. *American Psychologist*, 59, 7-13.

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

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440-452.

Schmader, T., Johns, M., & Barquissau, M. (in press). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*.

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

Sousa, P., Atran, S., & Medin, D. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, 2, 195-223.

Sousa, E., & Leyens, J-P. (1987). A priori vs. spontaneous models of attributions: The case of gender and achievement. *British Journal of Social Psychology*, 26, 281-292

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

Spielberger, C. D., Gorsuch, R. R., & Lushene, R. (1970). *The State-Trait Anxiety Inventory (STAI) test manual.* Palo Alto, CA: Consulting Psychologists Press.

Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology*, 75, 1191-1197.

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

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

Steele, C. M., & Aronson, J. (2004). Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist*, 59, 47-48.

Steele, C. M., Spencer, S. J., & Aronson, J., (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379-440). San Diego, CA, US: Academic Press, Inc.

Sternberg, R. J. (1985). *Beyond IQ*. New York: Cambridge University Press.

Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin*, 28, 1667-1678.

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

Way, N. (1998). *Everyday courage: The lives and stories of urban teenagers.* New York: New York University Press.

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*, 127, 797-826.

Wittenbrink, B., Hilton, J. L., & Gist, P. L. (1998). In search of similarity: Stereotypes as naive theories in social categorization. *Social Cognition*, 16, 31-55.

Yzerbyt, V. Y., Leyens, J-P., & Corneille, O. (1998). Social judgeability and the bogus pipeline: The role of naive theories on judgment in impression formation. *Social Cognition*, 16, 56-77.

Table 1: Scores on math 2 for "source of stereotype" and Setting conditions before and after covariation of Math 1 scores.

condition	n	mean	SD	Estimated marginal means[a]	SE of the estimated marginal mean[a]
G	28	3.57	2.06	3.73	.315
E	27	5.22	1.83	4.93	.323
ND	20	5.75	2.38	4.94	.393
Women only	34	4.65	2.21	4.85	.302
Mixed sex	41	4.34	2.15	4.22	.260

[a] Covariates appearing in the model are evaluated at the following values: Math 1 score=4.72

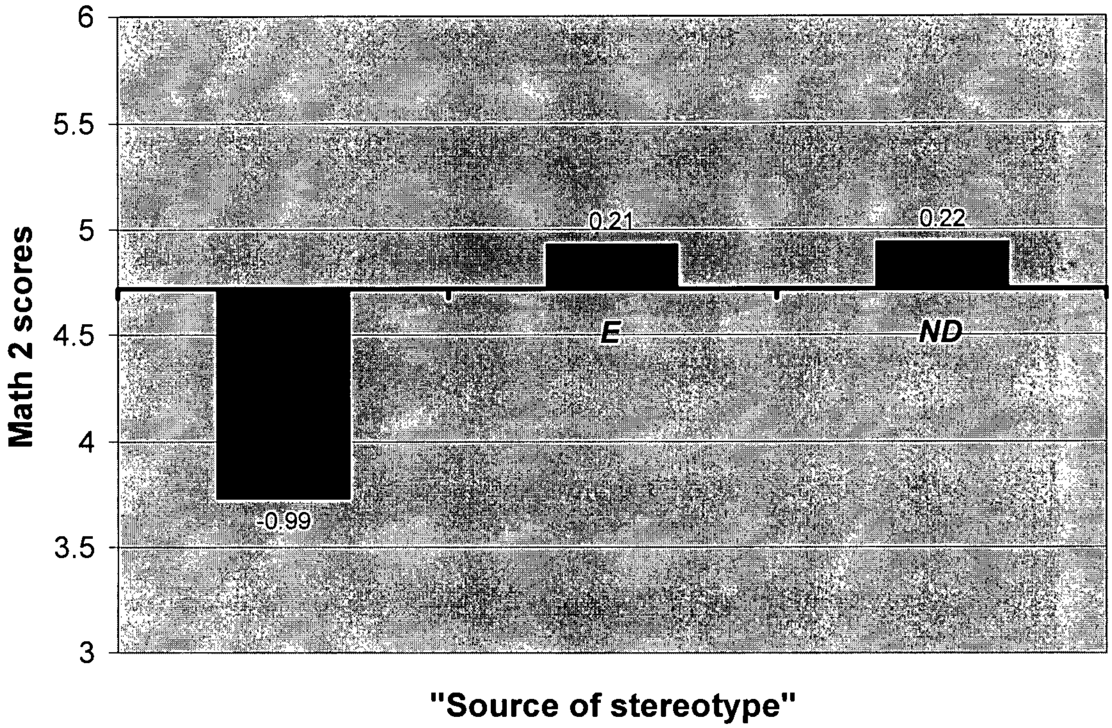
Figure Captions

Figure 1: Math 2 estimated means after controlling for Math 1 scores for each of the "source of stereotype" conditions

Figure 2: The relation between changes in effect size of "source of stereotype" conditions and accuracy

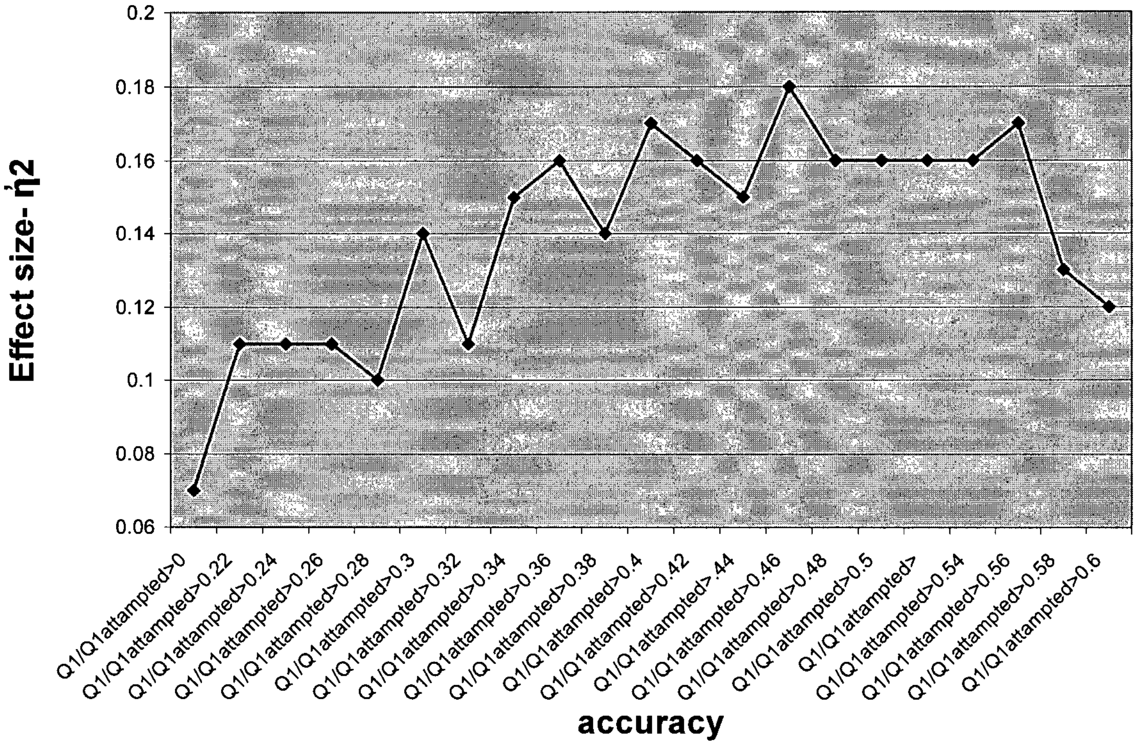
Figure 3: Jayaratne's (2002) suggested model for the relationship between genetic explanations and attitudes towards social groups.

Figure 1: Math 2 estimated means after controlling for Math 1 scores[a] for each of the "source of stereotype" conditions



[a] The Y axis origin was set at the level of the evaluated covariate (the mean for math 1 scores): Math 1 score=4.72.

Figure 2: The relation between changes in effect size of "source of stereotype" conditions and accuracy[a]



[a] The inclusion criterion for the central analyses was set at a chance level: $Q1_{score}/ Q1_{attempted}>.22$

Figure 3: Jayaratne's (2002) suggested model for the relationship between genetic explanations and attitudes towards social groups.

Genetic explanations for individual differences	→	Genetic explanations for perceived gender, class or race differences	→	Attitudes toward women, the poor or Blacks

Appendices

Appendix 1A: One version of the math sections

1. If $r \heartsuit s = r(r - s)$ for all integers r and s, then $4 \heartsuit (3 \heartsuit 5)$ equals

    a. -8
    b. -2
    c. 2
    d. 20
    e. 40

2. In the set of numbers (12, 5, 14, 12, 9, 15, 10), f equals the mean, g equals the median, h equals the mode, and j equals the range. Which of the following is true?

    a. $f > g > h > j$
    b. $g = h > f > j$
    c. $g > h > f = j$
    d. $j > f > g = h$

3. Which of the following CANNOT be an integer if the integer k is a multiple of 12 but not a multiple of 9?

    a. $\frac{k}{3}$
    b. $\frac{k}{4}$
    c. $\frac{k}{10}$
    d. $\frac{k}{12}$
    e. $\frac{k}{36}$

4. Shown below is a correct problem in multiplication, with z and y representing certain digits. What could the value of y be?

$$\begin{array}{r} zy \\ \underline{\times\ y} \\ 9y \end{array}$$

    a. 5
    b. 4
    c. 7
    d. 6
    e. 0

5. A rectangular bathroom measures 9½ feet by 12 feet. The floor is covered by rectangular tile measuring 1½ inches by 2 inches.

Column A	Column B
The number of tiles on the bathroom floor	6 x 48 x 19

a. the quantity in Column A is always greater
b. the quantity in Column B is always greater
c. the quantities are always equal
d. it cannot be determined from the information given

6. A professor is choosing students to attend a special seminar. She has 10 students to choose from, and only four may be chosen. How many different ways are there to make up the four students chosen for the seminar?

a. 24
b. 120
c. 180
d. 210
e. 360

7. k is an integer such that $9(3)^3 + 4 = k$

Column A	Column B
The average of the prime factors of k	16

a. the quantity in Column A is always greater
b. the quantity in Column B is always greater
c. the quantities are always equal
d. it cannot be determined from the information given

8. $a = b + 2$

Column A	Column B
$a^2 - 2ab + b^2$	$2a - 2b$

a. the quantity in Column A is always greater
b. the quantity in Column B is always greater
c. the quantities are always equal
d. it cannot be determined from the information given

9. If $m + n = 24$, and $m - n + p = 15$, then $4m + 2p =$

   a. 13
   b. 26
   c. 39
   d. 64
   e. 78

10. The distance between town A and town B is over 500 km. A car leaves town A at 10am going to town B at a speed of 70 km/h. A van leaves town B at 11:30am going to town A on the same road at a speed of 80 km/h. what is the distance between the vehicles 20 minutes before the rendezvous.

    a. 150 km
    b. 50 km
    c. 75 km
    d. 300 km
    e. there is not enough information to solve the problem

11. Sara, Jack, Mike, Katie, and Ron want to sit on a bench. How many different sitting arrangements are possible?

    a. 120
    b. 720
    c. 3125
    d. 60
    e. 90

12. How many solid square cubes with volume of 8 cm^3 can we fit in a rectangular box with the following dimensions: H-12 cm, W-9 cm, L-11 cm?

    a. 100
    b. 120
    c. 147
    d. 148
    e. 160

13. In a business meeting, there are 2 representatives from 12 different companies. How many handshakes will take place if everyone has to shake hands with all the people except the person from their own company?

    a. 552
    b. 529
    c. 276
    d. 264
    e. 248

14. If $3x + 4y$ is an odd number, which of the following CANNOT be true?
    a. x is odd and y is odd

b. x is even and y is even
c. x is even and y is odd

a. I only
b. I and II only
c. II only
d. III only
e. II and III only

15. What is the next number in the following series?

2, -1, -4, 11, 116, ___

a. 12,743
b. -7
c. 13,451
d. -11,227
e. -311

16. Tap A fills the pool in 20 hours Tap B fills the pool in 30 hours. Both taps were turned on at 8am. When will the pool be full?

a. 10pm
b. 8pm
c. 10am
d. 4am
e. 4pm

17. Jane loads a truck in 8 hours. Mandy unloads a truck in 10 hours. How long it will take to load an empty truck if both of them work together.

a. 40 hours
b. 18 hours
c. 26 hours
d. 80 hours
e. they will never get the work done

18. What is the perimeter of a rectangle that is three times as long as it is wide and has the same area as a circle of circumference of 6?

a. $8\sqrt{3\pi}$
b. $8\sqrt{\pi}$
c. $4\sqrt{3\pi}$
d. $\dfrac{8\sqrt{3}}{8}$
e. $\dfrac{8}{\sqrt{3\pi}}$

Appendix 1B: The second math section

1. $c \blacklozenge d = \dfrac{c-d}{c}$, *where* $c \neq 0$

   If $9 \blacklozenge 4 = 15 \blacklozenge k$, then $k =$.

   a. -3
   b. 6
   c. $\dfrac{20}{3}$
   d. $\dfrac{25}{3}$
   e. 9

2. If $x > 0$, then $(4^x)(8^x) =$

   a. 2^{9x}
   b. 2^{8x}
   c. 2^{6x}
   d. 2^{5x}
   e. 2^{4x}

3. The fuel economy of a passenger car will typically peak at 45 miles per hour, and then decline as the vehicle's speed increases. For example, a given car gets 40 miles per gallon at 45 miles per hour, but it gets 12% fewer miles per gallon at 60 miles per hour. How far can this car travel at 60 miles per hour on 11 gallons of gas?

   a. 35.2 miles
   b. 387.2 miles
   c. 400 miles
   d. 408.3 miles
   e. 440 miles

4. Simon and Chris have 26 plastic robots and 17 red toys. If they have a total of 38 toys, how many of the plastic robots are red?

   a. at least 17
   b. no more than 26
   c. no more than 7
   d. at least 5
   e. there is not enough information to solve the problem

5. What is the next number in the following series?

   5, 6, 11, 11, 23, 21, ____

   a. 45
   b. 22
   c. 47
   d. 39

e. 35

6. Judie started walking from her home to school at 9am. Kevin, her brother, noticed that she forgot her lunch and ran after her 30 minutes later. If Judie's pace is 4 km/h and Kevin's pace is 7 km/h, when will Kevin catch up with Judie?

   a. 10:30
   b. 10:10
   c. 10:00
   d. 10:50
   e. 10:15

7. If $m = 121 - 5k$ is divisible by 3, which of the following may be true?
   I. m is odd
   II. m is even
   III. k is divisible by 3

   a. I only
   b. II only
   c. II and III only
   d. I and II only
   e. I, II, and III

8. A box contains five blocks numbered 1, 2, 3, 4, and 5. Johnnie picks a block and replaces it. Lisa then picks a block. What is the probability that the sum of the numbers they picked is even?

   d. $\frac{9}{25}$

   e. $\frac{2}{5}$

   f. $\frac{1}{2}$

   g. $\frac{13}{25}$

   h. $\frac{3}{5}$

9. If A and B are positive integers and $24AB$ is a perfect square, then which of the following CANNOT be possible?
   I. Both A and B are odd
   II. AB is a perfect square
   III. Both A and B are divisible by 6

   a. I only
   b. II only
   c. III only

d. I and II only
e. I, II, and III

10. Shown below is a correct problem in addition, with x and y representing certain digits. What is the value of y?

$$\begin{array}{r} 7x \\ xy \\ +\quad xx \\ \hline 117 \end{array}$$

a. 1
b. 2
c. 3
d. 4
e. 5

11. If x and y are positive integers, and $x - 2y = 5$, which of the following could be the value of $x^2 - 4y^2$?

a. -3
b. 0
c. 14
d. 45
e. 51

12. If the result of squaring a number n is less than twice the number, then the value of n must be:

a. negative
b. positive
c. between -1 and +1
d. greater than 1
e. between 0 and 2

13. What is the perimeter of a rectangle that is twice as long as it is wide and has the same area as a circle of diameter 8?

a. $8\sqrt{\pi}$
b. $8\sqrt{2\pi}$
c. 8π
d. $12\sqrt{2\pi}$
e. 12π

14. Harvey paid $400 for a used car that travels 28 miles per gallon on the highway and 20 miles per gallon in the city. If he drove twice as many highway as city miles last month while using 34 gallons of gasoline, how many miles did he drive altogether.

a. 1,000
b. 840
c. 400

d. 340
e. 280

15. A plane is flying from City A to City B at m mph. Another plane flying from City B to City A travels 50 mph faster than the first plane. The cities are R miles apart. If both planes depart at the same time, in terms of R and m, how far are they from City A when they pass?

a. $\frac{R}{m}+50$
b. $\frac{Rm}{2m}-50$
c. $\frac{Rm}{2m+50}$
d. $\frac{R+50}{m+50}$
e. $\frac{m+50}{R}$

16. Ten real estate agents are meeting in a conference in Las Vegas. If they exchange business card at the beginning of the meeting so everyone gets a card from all the rest, how many business cards exchanged hands?

a. 200
b. 100
c. 90
d. 80
e. 110

17. In the set of positive, distinct integers $\{a, b, c, d, e\}$ the median is 16. What is the minimum value of $a + b + c + d + e$?

a. 26
b. 48
c. 54
d. 72

18. A fictitious credit card number always start with 4, has 8 numbers in total, and ends in an odd number. How many different combinations of numbers can make up this kind of credit card number?

a. 1,000,000
b. 10,000,000
c. 4,000,000
d. 5,000,000
e. 20,000,000

Appendix 2A: No gender differences in math abilities manipulation (ND condition essay)

Directions: After reading this passage, you will find a series of questions. Select the best choice for each question. Answers are based on the contents of the passage or what the author implies in the passage.

The New York Times
ON THE WEB

There are no gender differences in mathematical abilities, Researchers Say

By DR. ERIC A. GOODEY

The environmental camp in a longstanding controversial issue, which has drawn a lot of attention over the past few decades, has received the most convincing support to date in results released today from an international group of psychology researchers. The researchers claim to find that mathematical reasoning differences, which were assumed to exist between men and women, actually have no hold in reality. The results show that there are no innate differences between males and females in mathematical reasoning.

The new research is the largest published study of differences among males and females in mathematical reasoning. The research was conducted over 8 years in which the participants were followed and their performance closely observed. Unlike previous research in the field, the present study followed both a genetic research design (to look for inherit differences) and a survey research design. In the genetic paradigm, using top of the line instruments (F-MRI, DNA analyzers, and messenger RNA blockers), the researchers failed to find any gender differences on mathematical tasks.

Using the largest sample (over 50 million people!) ever recorded, the researchers compared grades in mathematics and physics in 113 countries around the world. The grades were taken from national comparison exams in elementary schools and high schools. The results showed that males and females were performing just as well on the math sections. The same results were found among college students in 26 countries.

The research was supported by the National Institute of Health (NIH), which provided the international team of researchers, led by Dr. Mark Goldstein from the Harvard Gender Research Institute, with a grant of an unprecedented 35 million dollars to fund a 6-year study of gender differences in education. The results that appeared today in Child Development, one of the leading journals of the American Psychology Association, are only the start of many that will follow in the coming years from this prolific team.

Dr. Thomas Schmidt, speaking for the team, concluded that "another bubble of prejudice had exploded. The sheer magnitude of the study is a guarantee of its results. If there would have been even the slightest difference between males and females in math, the study would have been able to detect it, up to a .001 difference in performance."

"This study is both statistically and clinically significant," said the leading author, Dr. Karen Dinear, director of child and adolescent psychiatry at the University of Wisconsin Medical Branch. "Its magnitude sheds new light on a long discourse concerning gender roles and math performance. I hope that teachers and student will modify their expectations accordingly."

Dr. Laura Wehr, from the University of North Dakota Microbiology and Genes Unit suggested that the results are not as sound as other may claim due to the size of the sample used in the genetic conditions of the study (950 females and 875 males). Other experts predicted more criticism in the coming weeks and months once more researchers in the field have a chance to review the findings.

1. What is the main argument of this article?

   a. Gender differences can only be detected by extremely sensitive tests
   b. A large sample population strengthens experiment results
   c. Males are not always better at math than females
   d. Gender differences are negated by social factors
   e. Mathematical skills are learned

2. How was mathematical reasoning represented in the experiment?

   a. Comparing algebra and problem solving questions
   b. Comparing mathematic grades
   c. Comparing mathematic and physics grades
   d. Comparing physics grades
   e. Comparing calculus grades

3. Why is this research the most convincing evidence to date?

   a. A large population was tested
   b. Use of advanced technological equipments is more reliable now
   c. Both innate and cognitive factors were tested
   d. Teachers opinions are more valued than others
   e. Experts conducted the experiment

4. Why are experts so excited about the experiment results?

   a. More funding will be provided
   b. Females will no longer shy away from mathematics
   c. Males and Females are equally capable of mathematical reasoning
   d. Gender differences can no longer be a good excuse for failing math class
   e. Many hours of hard work paid off

Appendix 2B: Males perform better on math tasks due to genetic disposition manipulation (G condition essay)

Directions: After reading this passage, you will find a series of questions. Select the best choice for each question. Answers are based on the contents of the passage or what the author implies in the passage.

The New York Times
ON THE WEB

Genes are involved in mathematical abilities, Researchers Say

By DR. ERIC A. GOODEY

The biological camp in a longstanding controversial issue, which has drawn a lot of attention over the past few decades, has received the most convincing support to date in results released today from an international group of genetic researchers. The researchers claim to find genetic bases for well-documented gender differences in mathematical reasoning abilities. The study shows that innate differences exist between males and females in mathematical reasoning.

The new research is the largest published study of polygenetic effects to test the interaction between different genes and higher cognitive functions. One of the main findings demonstrates an interaction of 2 genes located on the Y chromosome (which is found only in males) with genes on chromosome 5 and chromosome 7. This interaction produces hormonal changes guided by the hypothalamus. The onset of the hormonal release is guided by activation of the Brotically area in the frontal lobe. This area is activated when processing mathematical oriented problems. F-MRI scans show these hormonal changes create an increase in the amount of ATP (the body's currency of energy) molecules directed to the hippocampus when a person is engaged in higher mathematical reasoning tasks. The increased energy to this area of the brain, considered the "working memory organ", enables the person to retain more accessible short term memory information while concentrating, a critical element in mathematical reasoning capabilities. This genetic difference seems to explain the findings that boys show superior performance by having on average a grade 5 percentile points higher than girls.

The research was supported by the National Institute of Health (NIH), which provided the international team of researchers, led by Dr. Mark Goldstein from the Harvard Microbiology Research Institute, with a grant of an unprecedented 35 million dollars to fund a 6-year study of polygenetic effects on brain capacities. The results that appeared today in the Journal of American Medical Association are only the start of many that will follow in the coming years from this prolific team.

Dr. Thomas Schmidt, speaking for the team, concluded that "manipulating hormonal state in the same way that the polygenetic effect does, may enable us in the future to elevate females' mathematical reasoning abilities to be in-line with those of males".

"This study is both statistically and clinically significant," said the leading author, Dr. Karen Dinear, director of child and adolescent psychiatry at the University of Wisconsin Medical Branch. "Its magnitude sheds new light on a long discourse concerning the role that genes and the environment play in the finding that, in general, males have higher mathematical reasoning abilities than females."

Other experts said the study was important in adding to the limited knowledge about the effects of different hormones on brain functions.

Dr. Laura Wehr, from the University of Aiwa Microbiology and Genes Unit, suggested that the results are not as sound as other may claim due to the size of the sample used in the study (63 females and 58 males). Other experts predicted more criticism in the coming weeks and months once more researchers in the field have a chance to review the findings.

5. What is the main argument of this article?

   f. Males are better at math than females
   g. Females are better at math than males
   h. Males have a genetic math disadvantage over females
   i. Males have a genetic math advantage over females
   j. Males and Females both are genetically equipped for math

6. How does the "math gene" work?

   f. Through clearer visual representations
   g. Flow of energy allows longer short term memory retention
   h. Higher levels of cognitive thinking are encoded differently
   i. Hormones alter the structure of the brain
   j. More areas of the brain are triggered for enhanced mathematical attention

7. According to the article, what is not the cause of math differences between the sexes?

   f. Pituitary Gland
   g. Hormones
   h. Genes
   i. Hypothalamus
   j. Frontal lobe

8. According to this article, in the future how can females improve their math skills?

   a. Taking herbal supplements
   b. Asking more questions during math class
   c. It is not possible for females to improve their math skills
   d. Spending more time on their math homework
   e. Altering hormone secretions

Appendix 2C: Males perform better on math tasks due to experiential circumstances manipulation (E condition essay)

Directions: After reading this passage, you will find a series of questions. Select the best choice for each question. Answers are based on the contents of the passage or what the author implies in the passage.

The New York Times
ON THE WEB

Expectations are responsible for gender differences mathematical abilities, Researchers Say

By DR. ERIC A. GOODEY

The environmental camp in a longstanding controversial issue, which has drawn a lot of attention over the past few decades, has received the most convincing support to date in results released today from an international group of psychology researchers. The researchers claim to find reasons for well-documented gender differences in mathematical reasoning abilities. The results show that there are no innate differences between males and females in mathematical reasoning.

The new research is the largest published study of differences among males and females in mathematical reasoning. The research was conducted over 8 years in which the participants were followed and their performance closely observed. Unlike previous research in the field, the present study followed both a genetic research design (to look for internal factors to explain the difference) and a cognitive research design (to look for external factors to explain the difference). In the genetic paradigm, using top of the line instruments (F-MRI, DNA analyzers, and messenger RNA blockers), the researchers failed to find any gender differences on mathematical tasks.

Using an ingenious cognitive paradigm, the researchers manipulated the teachers' expectations of students in 64 elementary school classes in 18 cities and towns around the country. In the experimental condition, the researchers visited schools as educational psychologists and gave students a bogus mathematical test at the beginning of the year. Afterwards, they provided the teachers with fake reports that illustrated that the girls in the class were better in mathematics. Observing the teachers through a video camera in the class, it became apparent that teachers were paying more attention to the girls, were more praising towards them and were more dismissive of the boys. In the control condition, where no manipulations of teachers' expectations had taken place, the opposite pattern was observed. Teachers were more attentive towards boys, were praising them more and were more dismissive of girls. The findings showed that the girls in the experimental condition were superior to the boys if the teachers' expectations were manipulated in one of the first three years of elementary school followed by two more years. Manipulation of teachers' expectations after the third year seems to mitigate the effects of teacher's expectations in the first years of school, but not to enough to turn them around as in the case of a manipulation during the first three years. In the control conditions, boys showed superior performance by having on average a grade 5 percentile points higher than the girls throughout the 8 years of the experiment, providing more support to the general stereotype. The critical period for the students' self-expectations construct seems to be in the beginning of the formal education.

The research was supported by the National Institute of Health (NIH), which provided the international team of researchers, led by Dr. Mark Goldstein from the Harvard Gender Research Institute, with a grant of an unprecedented 35 million dollars to fund a 6-year study of gender differences in education. The results that appeared today in Child Development, one of the leading journals of the American Psychology Association, are only the start of many that will follow in the coming years from this prolific team.

Dr. Thomas Schmidt, speaking for the team, concluded that "manipulating teachers' expectations in the same way that the stereotype does, shows that the construct of mathematical abilities that is apparent in teachers' minds and behavior may as well be the factor that explains gender differences in math". The current research joins a long line of research showing the effect of teachers' expectations on students' performance.

"This study is both statistically and clinically significant," said the lead author, Dr. Karen Dinear, director of child and adolescent psychiatry at the University of Wisconsin Medical Branch. "Its magnitude sheds new light on a long discourse concerning the role that genes and the environment play in the finding that, in general, males have higher mathematical reasoning abilities than females."

Dr. Laura Wehr, from the University of North Dakota Microbiology and Genes Unit, suggested that the results are not as sound as other may claim due to the size of the sample used in the genetic conditions of the study (950 females and 875 males). Other experts predicted more criticism in the coming weeks and months once more researchers in the field have a chance to review the findings.

9. What is the main argument of this article?
   k. Mathematics should not be taught in co-ed classes
   l. Gender differences cannot be accounted for by innate qualities
   m. No reasonable explanations can account for differences in mathematical abilities
   n. Teachers should be aware of gender differences
   o. Girls are not putting enough effort into their math studies

10. Why is this research the most convincing evidence to date?
   k. A large population tested and the time spent on observation
   l. Use of advanced technological equipments is more reliable now
   m. Both innate and cognitive factors were tested
   n. Teachers opinions are more valued than others
   o. Children were unaware of the manipulation

11. According the article, how do math differences occur amongst boys & girls?
   k. Teachers' high expectations led girls to be more anxious and boys to be more determined
   l. Boys were disruptive affecting girls' concentration

m. Girls did not show as much interest in math as boys
n. Teachers were much more likely to help and praise boys than girls
o. Boys played with toys that involved more mathematical reasoning

12. In the future, how can females improve their math skills?

a. Take herbal supplements
b. Ask more questions during math class
c. Find teachers who praise them more
d. Teachers should be aware of their own interactions with students
e. Believe in their abilities