

# Gender-, race-, and income-based stereotype threat: the effects of multiple stigmatized aspects of identity on math performance and working memory function

Michele Tine · Rebecca Gotlieb

Received: 2 August 2012 / Accepted: 23 April 2013 / Published online: 9 May 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** This study compared the relative impact of gender-, race-, and income-based stereotype threat and examined if individuals with multiple stigmatized aspects of identity experience a larger stereotype threat effect on math performance and working memory function than people with one stigmatized aspect of identity. Seventy-one college students of the stigmatized and privileged gender, race, and income-level completed math and working memory pre-tests. Then, participants heard a moderately explicit stereotype threat-inducing prime. Next, participants took math and working memory post-tests. Stereotype threat effects were found on math performance on the basis of race and income-level, but not on the basis of gender. Stereotype threat effects were found on working memory function on the basis of gender, race, and income-level. For both measures, the income-based effects were the strongest. Results also suggest the possibility of multiple minority stereotype threat effects on math performance and working memory. More specifically, individuals with three stigmatized aspects of identity experienced significantly larger stereotype threat effects than those with zero-, one-, or two-stigmatized aspects of identity.

**Keywords** Stereotype threat · Math · Working memory · Identity · Cognition

## 1 Introduction

Systematic differences in academic achievement among varying demographic groups are a large and persistent source of concern in the U.S. educational system. Specifically, robust evidence supports a gender-based math achievement gap (Hyde et al. 1990; Leahey and Guo 2001; Muller 1998; Willingham and Cole 1997), a race-based math

---

M. Tine (✉) · R. Gotlieb  
Department of Education, Dartmouth College, 6103 Raven House, Hanover, NH 03755, USA  
e-mail: michele.tine@dartmouth.edu

achievement gap (Lubienski 2002; US Department of Education 2009), and an income-based math achievement gap (Brooks-Gunn et al. 1993; Lockheed et al. 1985; Palardy 2008; Yang 2003). These achievement gaps may be explained in part by the stereotype threat experienced by the underperforming groups. Stereotype threat occurs when negative stereotypes about a stigmatized group are made salient and people belonging to the stigmatized groups fear confirming these negative stereotypes, which, in turn, causes them to perform worse thereby perpetuating the stereotype (Steele and Aronson 1995). Gender-, race-, and income-based stereotype threat effects have all been independently documented. However, the literature has not yet directly investigated how these distinct forms of stereotype threat compare to one another. Furthermore, the literature has not fully examined how these distinct forms of stereotype threat operate in individuals that simultaneously experience more than one of the forms, even though such people clearly exist. The main purpose of the current study was to further our understanding of stereotype threat by (1) examining the relative impact of gender-, race-, and income-based stereotype threat effects and (2) determining how these stereotype threat effects become compounded in individuals that experience more than one of these stereotype threat effects. Understanding how gender-, race-, and income-based stereotype threat effects compare to one another and compound may be useful in focusing the most appropriate interventions on populations with the greatest needs. Further, it can powerfully inform stereotype threat research by suggesting demographic requirements for appropriate participant samples. Therefore, the current study aims to provide a more nuanced perspective of stereotype threat by uncovering if different types of stereotype threat operate in different ways. To more thoroughly understand the motivation of the study, it is important to review what the research has already uncovered about single- and multiple- minority stereotype threat effects.

### 1.1 Single minority stereotype threat effects on mathematics performance

*Single minority stereotype threat* occurs when someone identifies with one specific group, is aware of a stereotype about that group, and his or her behavior changes in a way that conforms to that group's stereotype. The most extensively studied example of single minority stereotype threat is *gender based-stereotype threat*. For example, when females are primed to be aware of their gender, they show a significant decline in their *mathematics performance*, or their ability to solve challenging mathematics test items (e.g., Davies et al. 2002; Good et al. 2008; Keller and Dauenheimer 2003; Schmader 2002; Shih et al. 1999; Steele et al. 2002; Walsh et al. 1999). Also supporting gender-based stereotype threat, Spencer et al. (1999) found that females underperformed compared to males on math tasks in which they believed females had previously demonstrated inferior performance. Even among females who self-select into pursuing a course of study in math, performance on a math test was shown to be susceptible to stereotype threat (Good et al. 2008).

*Race-based stereotype threat* exists as well. African Americans and Latino(a)s are stereotyped to perform worse than Caucasians in academic domains and many studies have shown that making African Americans or Latino(a)s aware of this stereotype causes them to perform worse on academic tasks than when they are not made aware

of these stereotypes (e.g., Steele and Aronson 1995; Brown and Day 2006; Gonzales et al. 2002; McFarland et al. 2003; Nguyen et al. 2003; Ployhart et al. 2003; Sawyer and Hollis-Sawyer 2005; Schmader and Johns 2003).

Finally, there is emerging evidence of *income-based stereotype threat*. People from families with a low income-level are stereotyped to have inferior intellectual abilities to those from families with a high-income level (Cozzarelli et al. 2001). When low-income students were told to take a test diagnostic of intelligence, they performed worse than high-income peers; however, when the test was not framed as diagnostic of intellectual ability, they performed the same as their high-income peers (Croizet and Claire 1998). In another study, the test performance of two groups of low-income students differed as a function of whether they were told the test they were to take was diagnostic of intelligence; the participants who had the test framed as diagnostic performed worse than the ones who did not (Spencer and Castano 2007). Moreover, when low-income college students were told that underprivileged students generally perform worse than other students, their performance on math and verbal tasks was significantly worse than when they were not told this information (Harrison et al. 2006). Similarly, Spencer and Castano (2007) found that asking low-income young adults to answer demographic questions about their parents' income and occupations before a test resulted in those participants performing more poorly than the low-income participants who were not asked these questions before the test. Taken together, the aforementioned studies provide ample evidence for gender-, race, and income-based single minority stereotype threat.

## 1.2 Multiple minority stereotype threat effects on mathematics performance

But importantly, many individuals belong to multiple stigmatized groups. These individuals face challenges distinct from the challenges an individual with only one stigmatized aspect of identity experiences (Brown and Leaper 2010; Purdie-Vaughns and Eibach 2008; Scott-Jones and Clark 1986). Specifically, some research suggests that these individuals experience *multiple minority stereotype threat* or that individuals that identify with multiple stigmatized aspects of identity experience a larger decrement to test performance than individuals that identify with only one stigmatized aspect of identity when under stereotype threat conditions. For example, Gonzales et al. (2002) found that ethnic minority females, who have two stigmatized aspects of identity, performed less well on cognitive tasks after being primed than did ethnic majority females, who have only one stigmatized aspect of identity. Gresky et al. (2005) findings also support the multiple minority stereotype threat phenomenon. They found that among stereotype primed females, performance on a math test was better for those who were instructed to think about the numerous parts of their identity (most of which were irrelevant to gender stereotypes and thus not affected by gender-based stereotype threat) than for those who were not given this instruction. Since awareness of aspects of identity irrelevant to a stereotype can decrease the effects of stereotype threat, it seems that making a person aware of multiple aspects of his or her identity, all of which are relevant to a stereotype, may make the effects of stereotype threat even greater.

Further, individuals who are members of two identity groups with conflicting stereotyped performance in the same domain (i.e., Asian females, who are stereotyped to be

bad at math because they are female, while also stereotyped to excel at math because they are Asian) have performed better than controls on a stereotype relevant task when the positive stereotype is primed (being Asian) and worse when the negative stereotype (being female) is primed (Shih et al. 1999; see also McGlone and Aronson 2006). This finding also lends credence to the idea that a multiple minority stereotype threat phenomenon may exist; if opposing stereotyped identities can cause opposite effects on performance, it seems possible that similarly stereotyped identities might have an additive effect on performance.

Similarly, Rosenthal and Crisp (2006) showed that women who were instructed to think about the similarities between men and women were less susceptible to gender-based stereotype threat; they performed better on a math task under stereotype threat conditions than did women who had not previously thought about the similarities between the sexes. In other words, thinking about the similarities between a stigmatized group and a non-stigmatized group decreased stereotype threat effects. It is plausible therefore, that thinking about the similarities between multiple different stigmatized groups to which an individual belongs could cause an additive stereotype threat effect, if all of the individual's stigmatized groups have a stereotype relevant to the task at hand. Finally, research findings suggest that priming a stereotype in more than one way can cause greater stereotype threat effects (Spencer and Castano 2007). Since making people more aware of a stereotype heightens stereotype threat effects, it seems that making people aware of more than one relevant stereotype might also make them more affected by stereotype threat. Notably, Harrison et al. (2006) found that inducing income-based stereotype threat did not affect people differently as a function of their race. Still it is possible that an additive multiple minority stereotype threat effect could exist if all those stigmatized aspects of identity are primed, rather than just one being primed. However, this hypothesis has never been explored.

### 1.3 Single minority stereotype threat effects on working memory

Recent work has also explored cognitive processes that are affected by stereotype threat. For example, there is initial evidence that stereotype threat may cause a decrease in *working memory capacity*, which is the limited amount of information one can maintain and process in his or her short term memory store at any given point in time (Baddedly and Hitch 1974; Rydell and Boucher 2010; Schmader and Johns 2003). For example, Schmader and Johns (2003) documented that stereotype threat reduced working memory capacity among members of stigmatized groups (see also Rydell and Boucher 2010). Since then, research has focused on the mechanisms that explain how and why working memory decreases under stereotype threat conditions. Schmader et al. (2008) suggest that stereotype threat leads to physiological arousal, performance monitoring, and negative thought suppression, which together affect working memory by taxing the central executive. Meanwhile, Beilock et al. (2006) found that stereotype threat disrupts the phonological loop, a sub-system of working memory that stores and rehearses verbal information, impairing performance of those who are threatened on math tasks. Additional evidence that working memory is affected by stereotype threat comes from Rydell et al. (2009) finding that making a female aware of multiple

aspects of her identity, some of which have positive associations with math performance, eliminated working memory decline caused by stereotype threat and improved performance on a math test. Forbes and Schmader (2010) found that retraining females to associate being female with possessing strength in math resulted in an increase in working memory capacity and a subsequent boost in performance on a math test under stereotype threat. Thus, when stereotype threat is decreased, there is a corresponding increase in working memory capacity.

Notably, some have argued that working memory reduction is not the mechanism responsible for stereotype threat effects (Beilock et al. 2007). For example, Jamieson and Harkins (2007) asserted that increased effort, and not a taxed working memory, explains the decrement in performance by the threatened group. As such, stereotype threat effects on working memory and effort will be examined in the current study, in addition to stereotype threat effects on mathematics performance.

#### 1.4 Multiple minority stereotype threat effects on working memory

While the effects of multiple minority stereotype threat on *mathematics performance* have been investigated preliminarily, its effects on working memory have not been explored. Based on the aforementioned body of research, it seems plausible that priming multiple stigmatized aspects of identity may result in a larger decrease in working memory than when only one stigmatized aspect of identity is primed. For example, Forbes and Schmader (2010) finding that decreasing stereotype threat increased working memory performance suggests that increasing stereotype threat (by priming multiple stigmatized aspects of identity) could decrease working memory performance.

#### 1.5 Research motivation, goals, and hypotheses

Thus, the goal of the present study is twofold. The first goal is to document single minority gender-, race-, and income-based stereotype threat effects on math performance and working memory function. Importantly, this is the first time income-based stereotype threat effects on working memory will be explored. Further, this goal affords us the opportunity to determine the relative impact of gender-, race-, and income-based standard stereotype threat effects, which has not previously been done. It is expected that the effect of gender-based stereotype threat will be weaker than the effect of race- or income-based stereotype threat. We have both empirical and conceptual justifications for this hypothesis. First, other research suggests race-based stereotype threat may be stronger than gender-based stereotype threat (Gonzales et al. 2002; Nguyen and Ryan 2008; Stricker and Ward 2004). Additionally, larger performance differences on mathematics tasks have been shown on the basis of race and income than on the basis of gender (Scott-Jones and Clark 1986). Conceptually, it seems possible that stereotype threat effects may impact stigmatized groups of smaller sizes in more extreme ways than stigmatized groups of larger sizes. For example, women comprise about half of our societies population, a fairly large group. Meanwhile, non-Caucasian and low-income individuals comprise smaller proportions of our societies population (U.S. Census 2009). An individual may feel that being associated with a smaller group

is more central to his identity than being part of a larger group, as being part of that smaller group is what makes him unique. In turn, these individuals may experience more severe stereotype threat effects than those that identify with larger identity groups. In the current study, this reasoning leads us to hypothesize that income- and race-based stereotype threat effects may be more extreme than gender-based effects, as stigmatized income- and race-groups are smaller in size than the stigmatized gender group.

The second goal of the present study is to determine whether there is a multiple minority stereotype threat effect on math performance and/or working memory function. Based on the literature above, it is hypothesized that there will be an additive multiple minority stereotype threat effect on both outcomes. That is, it is hypothesized that individuals with multiple stigmatized aspects of identity will experience more severe stereotype threat effects when primed than those who have fewer stigmatized aspects of identity, for the empirical and conceptual reasons previously outlined.

Ancillary analyses will also investigate if single and/or multiple minority stereotype threat effects impact the amount of effort individuals put forth on tasks. The hypotheses regarding the impact on effort remains open, as the literature reports contradictory findings. Some report that single minority stereotype threat increases effort (Beilock et al. 2006; Jamieson and Harkins 2007), others hypothesize that it decreases effort (Steele and Aronson 1995), and one concludes that there is no relationship between single minority stereotype threat and effort (Smith 2004).

## 2 Methods

### 2.1 Participants

Seventy-one undergraduate college students between ages 18 and 26 ( $M = 19.54$  year,  $SD = 1.602$ ) participated. Two different power analyses programs were used to determine an appropriate sample size (Power in Two-Level Designs and Optimal Design), each with unique flexibilities (Bosker et al. 2003; Raudenbush et al. 2011, respectively). Effect sizes, standard errors, variance–covariance matrices, variable means and standard deviations were based on prior research. Together, these two programs allowed consideration for the full range of planned analyses.

The mean quantitative SAT math reasoning test score was 694 ( $SD = 90.67$ ), and no participants scored below 450 out of 800.

#### 2.1.1 Stigmatized aspect of identity: gender

Twenty-five participants self identified as male. Forty-six participants self-identified as female and were considered to have a gender-based stigmatized aspect of identity (SAI).

#### 2.1.2 Stigmatized aspect of identity: race

In total there were 24 racial/ethnic minority (RM) participants; 17 of these identified as racially Black or African American. Seven participants identified as ethnically

**Table 1** Gender, race, and income-level of participants by NSAI group

NSAI group	N	Gender		Race		Income		
		Male	Female	White	RM	High-income	Middle-income	Low-income
0	20	20	0	20	0	16	4	0
1	27	3	24	25	2	18	8	1
2	14	2	12	2	12	4	6	4
3	10	0	10	0	10	0	0	10
Total	71	5	46	47	24	38	18	15

Hispanic or Latino(a) (and did not identify as racially White). These 24 RM participants were considered to have a race-based SAI. The other 47 participants identified as White.

### 2.1.3 Stigmatized aspect of identity: income

Participants were categorized as low-, middle-, or high-income based on an updated eight-income-bracket scheme used by [Harrison et al. \(2006\)](#). The income categories were adjusted upwards by \$6,000 to reflect the inflation that occurred in the six years that elapsed since that study was published. Therefore, participants from a family whose total annual income was <\$45,000 a year were categorized as low-income ( $n = 15$ ), participants from a family with an annual income between \$45,000 and \$84,999 a year were categorized as middle-income ( $n = 18$ ), and participants from a family earning over \$85,000 a year were categorized as high-income ( $n = 38$ ). The 15 participants who were low-income were categorized as having an income-based SAI, consistent with previous work that found stereotype threat effects on an academic task for low, but not middle or high-income participants ([Harrison et al. 2006](#)).

### 2.1.4 Number of stigmatized aspects of identity

The number of categories in which each participant was stigmatized was summed to create a measure of their total number of stigmatized aspects of identity (NSAI), ranging from zero to three. See [Table 1](#) for a breakdown of gender, race, and income of the zero-, one-, two-, and three-stigmatized aspects of identity groups.

## 2.2 Procedures

All participants were exposed to the same testing condition. When they came into the lab, a White female experimenter explained that they would take two tests, each divided into two sections to understand underlying cognitive factors involved in math problem solving. Then, participants were given ten minutes to work on the math pre-test. After that, participants completed the working memory pre-test. Next, all participants were primed with gender-, race-, and income-based stereotypes. Specifically, the experi-

menter read the following priming script based on previous stereotype threat prime scripts (see [Beilock et al. 2007](#); [Rydell et al. 2009](#)):

We are interested in investigating mathematical processing and particularly what makes some people perform better in math domains than others. As you likely know, math skills are important to performance in many subjects in college and quantitative reasoning skills predict long-term professional outcomes. Unfortunately however, there are systematic differences in math performance among people as a function of their race and socioeconomic background. There are also math performance differences between males and females. Surprisingly, little is known about the mental processes underlying math performance that might shed light on why these group differences occur. This research is aimed at better understanding why these race-, income-, and gender-based groups perform differently on math tasks.

Based on [Nguyen and Ryan \(2008\)](#) classifications of types of stereotype threat primes in their meta-analysis, this prime was “moderately explicit”, as it states that there are differences, but does not suggest which gender, race, or income level performs best. A moderately explicit prime was chosen because it was assumed that stereotypes about low-income people were more similar to stereotypes based on race than stereotypes based on gender. Both stereotypes about race and stereotypes about income level are about intelligence generally ([Cozzarelli et al. 2001](#); [Devine 1989](#)), whereas stereotypes about gender are often confined to the math domain ([Jacobs and Eccles 1985](#)). [Nguyen and Ryan \(2008\)](#) showed that racial minorities’ performance is more affected by moderately explicit primes than by subtle or blatant primes. Thus, a moderately explicit prime was thought to be most effective for inducing race- and income-based stereotype threat. Women’s performance is least affected by moderately explicit primes, but the difference in effect by type of prime has been shown to be smaller for gender than for race ([Nguyen and Ryan 2008](#)). Since this study was concerned with priming gender, race, and income-level stereotypes the moderately explicit prime was ideal.

After the prime, participants completed the math post-test. Next, they were primed again when the researcher said, “Now we are going to begin another task that will help us explore factors that may underlie race-, income-, and gender-based differences in performance on math tasks.” Then participants completed the working memory post-test. Finally, they completed an experiment experience and demographics survey, were thanked, debriefed, and compensated for their time.

## 2.3 Materials

### 2.3.1 Math pre- and post-test

The math pre- and post-tests each consisted of fifteen multiple-choice questions, each with five potential answers, from a paper and pencil based practice section of the quantitative Graduate Recording Exam (GRE). Participants had ten minutes to answer as many of the questions as they could. All questions were



word problems involving advanced algebra. The test, which was originally constructed by Schmader (2002) and used by Rydell and Boucher (2010) and Rydell et al. (2009), was divided into two sections of equal difficulty. Thus, participants took two ten-minute tests each with fifteen questions. Based on national performance norms of test takers who previously answered each question, the mean accuracy for the pre-test was 63.73% and the mean accuracy for the post-test was 63.80%. The mean number of questions that were attempted by participants (i.e., an answer was provided or scratch work was written indicating an attempted answer) on each test was approximately 11 (pre-test  $M = 10.97$ ,  $SD = 3.01$ ; post-test  $M = 11.04$ ,  $SD = 3.11$ ); the nationally normed mean accuracy for the first 11 questions was 62.5% for the pre-test and 64.1% for the post-test. All questions fell in a range of 44–80% accuracy among the national sample. The data of national norms came from Schmader's task; she provided the percent norms as reported in a GRE practice book. Schmader (2002) notes that these types of questions are appropriate for gender-based stereotype threat research because gender differences in math performance generally do not exist on tests of this level of difficulty.

### 2.3.2 Working memory pre- and post-test

The verbal working memory pre- and post-test was adapted from Alloway (2007) Automated Working Memory Assessment (AWMA) backward digit string lists and forward digit string lists. Participants were instructed to listen to a string of digits read to them from the computer while keeping their hands in their lap. For each digit string, they then were to recall the string in backwards order by writing the last digit of the string first, the middle digits in backwards order, and the first digit of the string last. For example, if they heard the string 3-4-9-1, they would write 1-9-4-3. Participants made no other marks on the page and erased nothing.

The digit strings began at a length of four digits and terminated at a length of nine digits. In total the task lasted nine and a half minutes. There were twenty seconds between the start of each digit string for string lengths four through seven and thirty seconds between the start of each digit string for strings lengths eight and nine. Time was not a constraint for participants. A digit was never repeated within a digit string. There were four trials at each digit string length. The only difference between the two versions of the working memory pre- and post-test was the actual digits.

### 2.3.3 Experiment experience, effort, and demographics survey

Participants were asked to complete a 21-item survey about their experiment experience, effort, and demographics. For example, one of the experiment experience items asked participants to report, using a 7-point Likert scale ranging from “strongly agree” to “strongly disagree” how much they “believe[d] that [their] task performance could confirm negative stereotypes about [their] race.” One of the effort items asked participants to “Please indicate how much effort you put forth on the math tasks you just completed” using a 7-point Likert scale that ranged from ‘extremely little’ to ‘extremely a lot’. The demographic questions asked participants to select, among the

options provided, their gender, race, ethnicity, age, family's total annual income, and SAT scores.

### 2.3.4 Scoring

The math pre- and post-tests were scored by calculating the number of questions correctly answered, which is consistent with the scoring scheme in several previous studies (Good et al. 2008; Johns et al. 2008; Spencer and Castano 2007; Steele and Aronson 1995). The highest possible score on each test was 15, which 3 participants achieved on the pre-test and 1 participant achieved on the post-test.

The scoring scheme for the working memory test was adapted from the AWMA scoring scheme. An answer on the working memory test was scored as correct if it was recalled in backwards order such that the first digit that participants heard was written last, the last digit first, and the intermediate digits were also in backwards order. If a participant correctly recalled two, three, or four digit strings out of a total of four strings at a particular digit string length, the participant's performance on the next level of digit strings would be considered. If a participant correctly recalled one or no digit strings at a particular digit string length their score terminated at that level. The overall score was the total number of correct digit strings recalled up to and including the digit strings at the level of termination. The highest possible level of termination was ten (i.e. that participants had yet to terminate at level 9, which was the level at which the test ended), which 9 participants achieved on the pre-test and 7 achieved on the post-test. The highest possible score on the working memory test was 24, which no participants achieved.

Survey responses about experiment experience and demographics were recorded. The free response question about what the research was studying was coded in binary as, "reported the stereotype relevant parts of the experiment," or "did not report the stereotype relevant parts of the experiment."

## 3 Results

### 3.1 Preliminary analyses

All participants, except for two who left the question blank, reported in an open-ended question that they believed the experiment was interested in how demographic factors affect performance, suggesting that the experimental manipulation worked. The two participants who left the open-ended question blank were still included in all subsequent analyses, as they were exposed to the same content as all of the other participants and appeared to be fully engaged in the experiment (i.e., they provided answers for every other item in the study.) The overall mean number of questions answered correctly out of 15 on the pre-test of math ability was 7.93 ( $SD = 3.83$ ) and 8.07 ( $SD = 3.61$ ) on the post-test of math ability. The mean score out of 24 on the working memory pre-test was 11.92 ( $SD = 5.18$ ) and 14.45 ( $SD = 4.87$ ) on the post-test. The overall pre- and post-working memory termination levels were 7.45 ( $SD = 1.63$ ) and 7.92 ( $SD = 1.32$ ), respectively.

There was a significant difference in math SAT scores as a function of number of stigmatized aspects of identity,  $F(3, 67) = 17.66, p < 0.001$ . The zero-SAI group ( $M = 753$ ) and the one-SAI group ( $M = 721$ ) scored significantly higher than the two-SAI group ( $M = 620$ ) and the three-SAI group ( $M = 601$ ). There was a significant difference in SAT score as a function of gender, race, and income,  $F(1, 69) = 12.21, p < 0.01$ ;  $F(1, 69) = 32.76, p < 0.001$ ;  $F(2, 68) = 16.97, p < 0.001$ . In each case the stigmatized group had a lower mean score. Females ( $M = 668$ ) scored lower than males ( $M = 741$ ) and RM participants ( $M = 622$ ) scored lower than White participants ( $M = 730$ ). High-income participants ( $M = 739$ ) scored significantly higher than middle- ( $M = 661$ ) or low-income ( $M = 615$ ) participants. These differences are not controlled for or covaried in future analyses because the use of SAT scores in a stereotype threat ANCOVA would directly violate a number of the assumptions of the theory of ANCOVA (Sackett et al. 2004; Wicherts 2005). Further, by virtue of the fact that all these students were admitted to and attend the same elite private college, even with differences in SAT score (which may be the result of differences in preparation for the test) it is likely that all participants have similar academic prowess.

### 3.2 Identity-based stereotype threat effects

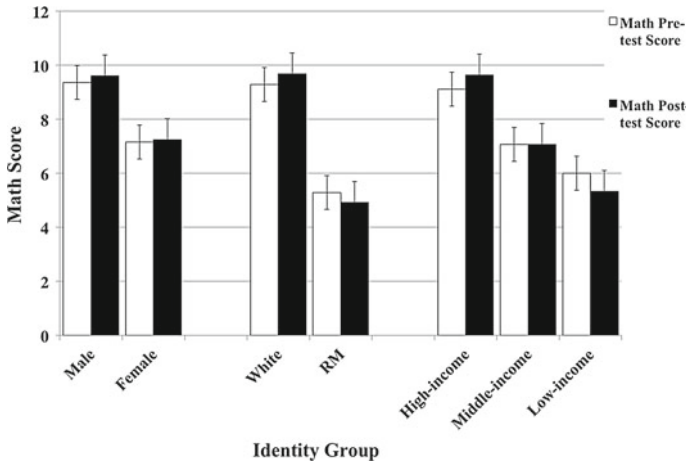
We tested for gender-based stereotype threat, race-based stereotype threat, income-based stereotype threat, and finally a multiple minority stereotype threat effect on the math test and on the working memory measure. In order to test for the effects of stereotype threat, we conducted ANCOVAs with post-test scores as the dependent measure and pretest scores as a covariate.

#### 3.2.1 Gender-based stereotype threat

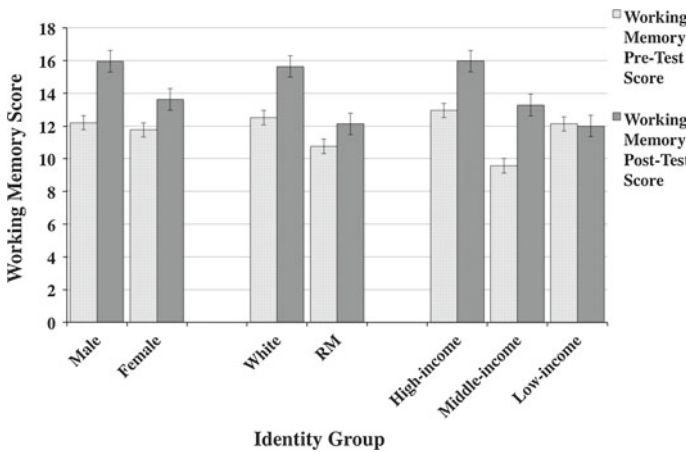
Results from an ANCOVA with gender as a between-subject variable and math pre-test score as a covariate revealed no stereotype threat effect on math test performance [ $F(1, 68) = 1.83, p > 0.05, \eta_p^2 = 0.026$ ] (see Fig. 1). There was, however, a significant effect of stereotype threat on working memory performance as revealed by an ANCOVA with working memory pre-test score as the covariate [ $F(1, 68) = 4.91, p < 0.05, \eta_p^2 = 0.067$ ]. Both females and males had higher performance on the post-test than the pre-test, but males ( $M_{\text{change}} = 3.76, p < 0.05$ ) improved more than females ( $M_{\text{change}} = 1.87, p < 0.05$ ) (see Fig. 2). An independent-samples t-test of self-reported endorsement of the statement, “My performance could confirm negative gender stereotypes about my gender” showed no significant difference between males and females ( $p > 0.05$ ). The fact that participants did not differ in this belief suggests that the females were not consciously experiencing stereotype threat effects.

#### 3.2.2 Race-based stereotype threat

Results from the ANCOVA with race as a between-subject variable and math pre-test score as a covariate revealed a significant stereotype threat effect on math test performance [ $F(1, 68) = 16.73, p < 0.001, \eta_p^2 = 0.197$ ]. White students slightly improved



**Fig. 1** Mean math pre- and post-test scores as a function of gender, race, and income-level group membership



**Fig. 2** Mean working memory scores as a function of gender, race, and income-level group membership

from pretest to post-test ( $M$  change = 0.40,  $p > 0.05$ ) whereas racial/ethnic minority students' performance slightly decreased ( $M$  change =  $-0.38$ ,  $p > 0.05$ ) (see Fig. 1). An ANCOVA with working memory pre-test scores as the covariate showed that there was also a significant stereotype threat effect on working memory performance [ $F(1, 68) = 7.41$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.098$ ]. All students had higher performance on the post-test than the pre-test, but White students ( $M$  change = 3.13,  $p < 0.05$ ) improved more than RM students ( $M$  change = 1.38,  $p < 0.05$ ) (see Fig. 2). Differences in self-report of stereotype threat experienced on the basis of race are not reported because White participants were not asked this question.

### 3.2.3 Income-based stereotype threat

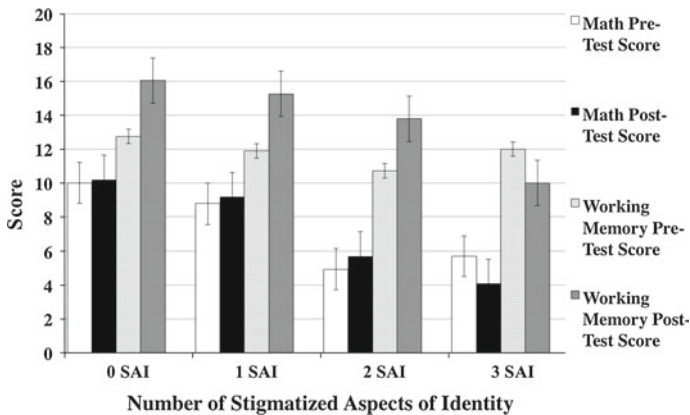
Results from the overall ANCOVA with income as a between-subject variable and math pre-test controlled for showed that there was a significant effect of income on math test performance [ $F(12, 67) = 5.92, p < 0.01, \eta_p^2 = 0.150$ ]. LSD post hoc tests revealed a significant difference between low-income participants and high-income participants ( $p < 0.01$ ) but there was no difference between low-income participants and middle-income participants ( $p > 0.05$ ) or middle-income participants and high-income participants ( $p > 0.05$ ). The low-income participants did more poorly at post-test than pre-test ( $M$  change =  $-0.67, p < 0.05$ ), whereas high-income participants improved ( $M$  change =  $0.53, p < 0.05$ ), and middle-income participants showed no change (see Fig. 1). This pattern is consistent with a stereotype threat effect among low-income participants relative to high-income participants but not relative to middle-income participants. This pattern also suggests that there may be a linear relation between income level and amount of stereotype threat experienced.

Results from the overall ANCOVA with income as a between-subject variable, working memory pre-test score as the covariate, and working memory post-test as the outcome showed that there was a significant effect of income [ $F(2, 67) = 4.92, p < 0.05, \eta_p^2 = 0.128$ ]. LSD post hoc comparisons show that the mean for the low-income group was significantly lower than the score for the middle-income group ( $p < 0.05$ ) and the high-income group ( $p < 0.01$ ) and that the middle- and high-income groups were not significantly different from one another ( $p > 0.05$ ). The low-income students performed slightly worse at post-test compared to pre-test ( $M$  change =  $-0.13, p > 0.05$ ), whereas middle-income participants ( $M$  change =  $3.72, p < 0.05$ ) and high-income participants ( $M$  change =  $3.03, p < 0.05$ ) both improved quite a bit from pre-test to post-test (see Fig. 2). This pattern is consistent with a stereotype threat effect among low-income participants relative to both middle- and high-income participants.

A one-way between-subject ANOVA of self-reported income-based stereotype threat experienced during the experiment as a function of one's income was also run. There was a significant effect of income on self-report of stereotype threat experienced for the three income levels [ $F(2, 68) = 6.15, p < 0.01$ ]. Post hoc comparisons using the LSD test indicate that the mean score for the low-income condition ( $M = 4.47, SD = 2.17$ ) was significantly higher than the mean score for the high-income condition ( $M = 2.68, SD = 1.69$ ), indicating greater feelings of threat. The middle-income mean score ( $M = 3.67, SD = 1.41$ ), however, was not significantly different from either the high or low-income group. Taken together these results indicate that low-income participants reported experiencing stereotype threat effects and that the extent to which participants reported experiencing income-based stereotype threat may be related to their income level.

### 3.2.4 Multiple minority stereotype threat

Results from an ANCOVA showed that there was a significant stereotype threat effect on math test performance as a function of the number of stigmatized aspects of identity (NSAI) that participants possessed when math pre-test scores were covar-



**Fig. 3** Mean math and working memory scores as a function of number of stigmatized aspects of identity

ied [ $F(3, 66) = 6.46, p < 0.01, \eta_p^2 = 0.191$ ]. A post hoc analysis showed that the effect was driven by the three-SAI group showing a significantly different change in performance than the zero-SAI group ( $p < 0.001$ ), the one-SAI group ( $p < 0.001$ ), and the two-SAI group ( $p < 0.05$ ), whereas the zero, one, and two groups did not differ from one another ( $p$ 's  $> 0.05$ ). The three-SAI group performed more poorly on the post-test than on the pre-test ( $M$  change =  $-1.60, p < 0.05$ ) while the zero, one, and two SAI groups all improved slightly on the post-test ( $M$  change =  $0.20, p > 0.05$ ;  $M$  change =  $0.41, p < 0.05$ ;  $M$  change =  $0.79, p < 0.05$ , respectively) (see Fig. 3).

There was also a significant stereotype threat effect on working memory performance [ $F(3, 66) = 6.82, p < 0.001, \eta_p^2 = 0.227$ ]. A post hoc analysis showed that the effect was again driven by the three-SAI group showing a significantly different change in performance compared to the zero-SAI group ( $p < 0.001$ ), the one-SAI group ( $p < 0.001$ ), and the two-SAI group ( $p < 0.01$ ), whereas the zero, one, and two groups did not differ from one another ( $p$ 's  $> 0.05$ ). The students in the three-SAI group performed more poorly on the post-test than on the pre-test ( $M$  change =  $-2.00, p < 0.05$ ), while the zero, one, and two SAI groups all improved on the post test ( $M$  change =  $3.30, p < 0.05$ ;  $M$  change =  $3.37, p < 0.05$ ;  $M$  change =  $3.07, p < 0.05$  respectively) (see Fig. 3).

### 3.3 Ancillary analysis

#### 3.3.1 Effort

The overall mean effort reported on a 7-point Likert scale ranging from 'extremely little effort' to 'extremely a lot of effort' was 5.65 ( $SD = 0.81$ ), indicating that students generally exhibited effort on the tasks. Independent-samples t-tests revealed no difference in self-reported effort between males and females or between White and RM participants,  $p > 0.05$  in both cases. A one-way between-subject ANOVA

revealed a significant difference between the reported effort of high-, middle-, and low-income participants [ $F(2, 68) = 3.18, p < 0.05$ ]. Post hoc LSD comparisons show that low-income participants reported exerting significantly more effort ( $M = 6.10, SD = 0.54$ ) than high- ( $M = 5.54, SD = 0.83$ ) or middle-income participants ( $M = 5.50, SD = 0.86$ ), but there was no significant difference between the self-report of effort exerted between high- and middle-income participants.

Although among the three demographic categories there was only a difference in effort on the basis of income (and not gender or race), there was an overall difference in the amount of effort as a function of the NSAI. A one-way between-subject ANOVA revealed a significant difference in effort among the four different NSAI groups [ $F(3, 67) = 3.50, p < 0.05$ ]. Post hoc comparisons using the LSD test showed that groups with more than one stigmatized aspect of identity between them were significantly different from one another in the effort they reported putting forth. Specifically, the three SAI group put forth more effort ( $M = 6.15, SD = 0.47$ ) than the zero ( $M = 5.30, SD = 0.92, p < 0.01$ ) or one ( $M = 5.57, SD = 0.63, p < 0.05$ ) SAI group but did not report exerting more effort than the two SAI group ( $M = 5.93, SD = 0.92, p > 0.05$ ). Similarly, the two stigmatized aspects of identity group put forth more effort than the zero stigmatized aspects of identity group ( $p < 0.05$ ), but not significantly more than the one or three SAI groups ( $p$ 's  $> 0.05$ ).

## 4 Discussion

### 4.1 Single minority stereotype threat

The first goal of the study was to document standard single minority gender-, race-, and income-based stereotype threat effects on math performance and working memory function.

#### 4.1.1 Gender-based stereotype threat

Surprisingly, there was no evidence of gender-based stereotype threat effects on math performance. The female participants did not do any worse on the math tasks after hearing a prime intended to induce gender-based stereotype threat. There were gender-based stereotype threat effects on working memory such that women improved less on the working memory post-test than did men.

Although it was hypothesized that there would be gender-based stereotype threat effects on math performance and working memory, there are several logical reasons why the stereotype threat effects on math were not observed. First, it has been suggested that the gender-based stereotype threat literature has overstated the scale of the phenomenon in multiple ways, such as the predominant use of ANCOVA with SAT scores as a covariate which violates ANCOVA assumptions and the use of participant matching paradigms (Sackett et al. 2004; Wax 2009; Wicherts 2005). While many studies do support the existence of gender-based stereotype threat effects, its prevalence has been exaggerated by a bias towards publishing studies that support stereotype threat and not publishing the many studies that have not found significant stereotype

threat results (Quayle 2011). Even among published studies, Stoet and Geary (2012) suggest that 9 out of 20 studies that attempted to replicate the original gender-based stereotype threat study failed to do so. All of this suggests that, while gender-based stereotype threat does likely exist, the magnitude and prevalence of this phenomenon has likely been overstated and there have been many cases, like the current study, where gender-based stereotype threat effects have not been observed. Second, the fact that the experimenter in this study was female may have lowered the gender-based threat participants felt. Third, a moderately explicit stereotype prime was used in this study, which has been shown to be the least effective type of prime for inducing gender-based stereotype threat (Nguyen and Ryan 2008). Indeed, Kray et al. (2001) found that, while subtle primes can cause people to perform in gender-stereotype consistent ways, blatant primes can cause people to act in a way that is gender-stereotype-inconsistent. Their explicit prime actually made females outperform males on the stereotype relevant task (Kray et al. 2001). It is possible that the present study produced a similar effect in female participants.

Interestingly, gender-based stereotype threat effects were observed on working memory. Specifically, both male and female participants performed better on the working memory post-test than the pre-test; however, male participants improved more than the female participants. Previous research has often demonstrated stereotype threat effects by a decrement in performance by the stigmatized group after hearing the stereotype threat prime. Although the current study did not find such a decrement, the present pattern of results still suggests that the females were experiencing stereotype threat since they improved less than the males. Perhaps practice on the working memory test caused improvement on the second administration for most, but stereotype threat stifled females' improvement. The finding of gender-based stereotype threat effects on working memory is consistent with previous literature and the hypothesized results (Beilock et al. 2007; Rydell and Boucher 2010; Schmader and Johns 2003).

It may be the case that gender-based stereotype threat effects were observable on working memory but not on math because stereotype threat affects these two types of processes differently. Although some previous research suggests that working memory mediates the relationship between stereotype threat and performance on cognitive tasks, the present research suggests that stereotype threat may affect working memory and math performance via different mechanisms (Beilock et al. 2007; Rydell and Boucher 2010; Rydell et al. 2009; Schmader and Johns 2003). Consistent with the present study, previous work has shown that stereotype threat effects can occur on tasks that do not require any working memory resources (i.e., golf putting) and thus that stereotype threat may result in decrements on different tasks for different reasons (Beilock et al. 2006). Jamieson and Harkins (2007) found that working memory interference does not explain stereotype threat effects on performance. Thus, the presence of gender-based stereotype threat effects on working memory but not math performance in the present research adds to the emerging literature that questions working memory as a mediator of stereotype threat effects on performance. This suggests that in assessing the impact of stereotype threat on individuals, it is important to consider not only the impact on performance measures, which are more commonly assessed, but also the way in which stereotypes may affect less visible cognitive processes.



#### 4.1.2 Race-based stereotype threat

Race-based stereotype threat effects on math performance and working memory were found. That is, hearing a prime intended to produce race-based stereotype threat effects did indeed make RM participants perform more poorly than White participants. The existence of race-based stereotype threat effects on math performance was consistent with expectations, as it has been observed in previous research (McFarland et al. 2003; Nguyen et al. 2003; Ployhart et al. 2003). Similarly, stereotype threat effects on working memory have been observed among racial/ethnic minorities (Schmader and Johns 2003). That the experimenter in the current study was of the threatening race, and that the type of prime used was the kind that has been shown to cause the greatest race-based stereotype threat effects, may have contributed to race-based stereotype threat effects emerging (Nguyen and Ryan 2008).

#### 4.1.3 Income-based stereotype threat

Consistent with expectations and previous research, there was evidence of income-based stereotype threat effects on math performance (Harrison et al. 2006). On the math task, low-income participants performed significantly worse on the post-test (with pre-test scores as a covariate) than did the high-income participants. In fact, low-income participants did worse on the math post-test than they had on the pre-test, while high-income participants improved and middle-income participants stayed the same.

Low-income participants also experienced stereotype threat effects on working memory. They performed worse on the working memory post-test than on the pre-test, whereas high- and middle-income participants all *improved* on the working memory post-test, likely due to a practice effect. Importantly, this is the first research to document the existence of income-based stereotype threat effects on working memory. That is, priming income-based differences adversely affects low-income people's working memory performance, but not middle- or high-income people's working memory performance. This is consistent with previous research about gender- and race-based stereotype threat effects on working memory and research about income-based stereotype threat effects on academic performance (Beilock et al. 2007; Harrison et al. 2006; Rydell and Boucher 2010; Schmader and Johns 2003). Self-report of stereotype threat experienced in the experiment also corroborated that low-income participants experienced stereotype threat relative to high-income participants.

## 4.2 Multiple minority stereotype threat

The second goal of the study was to determine if there are multiple minority stereotype threat effects on math performance and working memory function; the results provide some evidence for multiple minority stereotype threat for both outcomes. That is, participants in the three-SAI group experienced stereotype threat effects on math performance, while participants in the zero-, one-, and two-SAI groups did not. The math performance of participants with three stigmatized aspects of identity *decreased*

after hearing the stereotype-relevant primes, whereas math performance of all other participants *improved* slightly on the post-test. The same pattern was found on the working memory measures. That is, participants with three-SAIs exhibited a drop in working memory function after the prime, whereas participants in the zero, one and two-SAI groups did not experience stereotype threat effects and in fact showed a boost in performance on the working memory post-test.

However, it is important to be cautious about interpreting the results found in the current study as irrefutable evidence of multiple minority stereotype threat effects. We say this for two reasons. First, the nature of the effect was slightly different than expected. Specifically, these results suggest that there is not an *additive* effect as a function of additional stigmatized aspects of identity. Were there an additive effect of NSAI, there would be a step-wise increase in stereotype threat such that the zero-SAI group would not show stereotype threat effects, the one-SAI group would show small stereotype threat effects, the two-SAI group would show greater stereotype threat effects, and the three-SAI group would show the greatest stereotype threat effects. However, this pattern did not hold. Rather, the multiple minority stereotype threat only seemed to impact performance of those participants with three stigmatized aspects of identity. It is not surprising that a clean additive effect was not present based on the finding that there was not a gender-based stereotype threat on math performance. In turn, gender was not able to contribute to a step-wise additive multiple minority stereotype threat effect. There were race- and income-based stereotype threat effects, but these did not show clear, additive, multiple minority stereotype threat effects either. It is possible that because being a racial minority and low-income are closely related in this country ([United States Census 2010](#)), the stereotype threat experienced on the basis of just race, or just income, induced the stereotype threat that would be experienced on the basis of just the other category as well. That is, perhaps, when an individual experiences race-based standard stereotype threat effects, the similarity between race and income in this country induces them to simultaneously experience income-based threat as well. [Harrison et al. \(2006\)](#) work suggests that this might not be the case; however, they did not examine the extent to which individuals see race and income as related. For individuals who do see these aspects of identity as closely tied, stereotypes about race or income could induce stereotype threat effects for the other as well.

The second reason that we are careful when interpreting the results for multiple minority stereotype threat, is that it is possible the three-SAI group effects are in part a reflection of the strong income-based stereotype threat effect, as opposed to a compounded effect of three stigmatized aspects of identity. In other words, one reason the three-SAI group experienced greater stereotype threat than did the one- or two-SAI group is that the three-SAI group was composed of a higher proportion of individuals experiencing income-based stereotype threat. Specifically, while every participant in the three-SAI group was low-income, only 29 % (4/14) of the participants in the two-SAI group were low-income, and 4 % (1/27) of the participants in the one-SAI group were low-income. For these reasons, we interpret the multiple minority stereotype threat effects with some precaution. Future work should be specifically designed to unpack single minority income based stereotype threat effects from multiple minority stereotype threat effects by having more heterogeneity in the types of SAIs in the one-, two-, and three-SAI groups.

It is also possible that the income-based stereotype threat effects emerged as the most robust type of stereotype threat effects in this experiment because there was a larger proportion of people with three-SAIs experiencing income-based stereotype threat than there was experiencing gender- or race-based stereotype threat. For example, in the sample, 67 % (10/15) of participants who experienced income-based stereotype threat had three-SAIs, about 42 % (10/24) of participants who experienced race-based stereotype threat had three-SAIs, and 22 % (10/46) of participants who experienced gender-based stereotype threat had three-SAIs. Similarly, perhaps gender-based stereotype threat effects were the least robust because the smallest proportion of people who had three-SAIs was in the gender-based stereotype threat condition as compared to the race- and income-based stereotype threat conditions. If people with multiple stigmatized aspects of identity are the individuals primarily responsible for the appearance of stereotype threat effects on the basis of a single aspect of identity, then stereotype threat researchers need to be particularly sensitive to the multiple aspects of identity that their participants possess and not just the aspect of identity that is of interest to the researcher. For example, if researchers are examining gender-based stereotype threat, they should ensure that their participants possess only a stigmatized gender identity and not a stigmatized race- or income-based identity. This would ensure that any observed effects are actually gender-based and not the result of threat felt by an individual with multiple SAIs. This is especially critical for stereotype threat research that uses subtle primes that are not specific to one aspect of identity (e.g., “this test is diagnostic of ability”), as these sorts of primes would likely have a greater effect on participants with multiple SAIs than participants with only the one relevant SAI. This follows from [Gonzales et al. \(2002\)](#) finding that using a prime about test diagnosticity did indeed have a greater effect on participants with more than one SAI than it did on participants with only one or no SAIs. Related, standardized tests that claim to be diagnostic of intelligence or ability may be producing stereotype threat effects, particularly for individuals with multiple SAIs. This could be contributing to achievement gaps.

### 4.3 Comparing single minority stereotype threat effects

One of the most interesting aspects of the present study is that it affords a comparison of the relative impact of gender-, race-, and income-based stereotype threat effects. Income-based stereotype threat affects participants significantly more than do gender- or race-based stereotype threats. There were no gender-based stereotype threat effects on math performance, but there were clear income-based stereotype threat effects. While gender- and race-based stereotype threat effects on working memory were shown by a smaller increase in working memory scores among the stigmatized group than the non-stigmatized group, income-based stereotype threat effects on working memory were caused by the stigmatized group's *decrement* in performance. Further, a comparison of the effect sizes of the three types of stereotype threat reveals that income-based stereotype threat had a much more sizable impact than the other two. Finally, the composition of the SAI groups (as described above), such that the three-SAI group experienced income-based stereotype threat effects, while most participants in the other three groups were not experiencing income-based stereotype threat, may

explain the pattern of results observed as a function of NSAI. Perhaps income-based stereotype threat is what drove the three-SAI group to exhibit stereotype threat effects while the other groups did not. This too, would indicate the relative strength of income-based stereotype threat effects.

One reason why income-based stereotype threat effects were the most pronounced among the three types of stereotype threat may be that by now students are somewhat familiar with the idea of gender- and race-based stereotype threat. Due to the relative lack of income-based stereotype threat research compared to gender- or race-based stereotype threat research, the lack of attention to issues of income-level in research on stigmatization more generally, and the discomfort people experience in talking about issues of income level, people may not yet be familiar with the idea of income-based stereotype threat effects (Williams 2009). Thus, participants may have been more susceptible to income-based than to gender- or race-based stereotype threat because of a lack of knowledge about income-based stereotype threat.

#### 4.4 Effort

There was no difference between males and females or between White and RM participants in effort reported on the tasks. Low-income participants reported exerting significantly more effort on the tasks than did high- or middle-income participants. There was a significant difference in effort exerted as a function of the number of stigmatized aspects of identity, such that groups that had more than one stigmatized aspect of identity separating them differed in the amount of effort they exerted on the task. For example, the three-SAI group exerted significantly more effort than the one- or zero-SAI group but not than the two-SAI group.

To date there have been mixed results about the impact of stereotype threat on effort exerted by participants. Consistent with this study, some previous research has indeed suggested that participants under stereotype threat exert more effort than those who are not under stereotype threat (Beilock et al. 2006; Jamieson and Harkins 2007). However, Smith (2004) reviewed 18 studies that operationalized effort in three different ways (two behavioral measures and self-report) and found that effort did not wholly or even partially explain stereotype threat effects. In light of the present finding that participants needed to have more than one stigmatized aspect of identity between them in order for stereotype threat effects to be observed, perhaps much of the previous literature has failed to find stereotype threat effects on effort because it was examining the difference in effort exerted among participants primed with only one stigmatized aspect of identity (e.g. Keller 2002; Keller and Dauenheimer 2003).

#### 4.5 Limitations

It is important to note some limitations of the current study. For example, and as stated earlier, the distinct composition of the NSAI groups did not make for an ideal comparison of the effects of NSAI. To some extent the one-SAI, two-SAI, and three-SAI groups actually compared gender-based stereotype threat, to gender- and race-based stereotype threat, to gender-, race-, and income-based stereotype threat, respectively.

Also, the sample in the current study was comprised of undergraduate students from a highly selective institution. The way that the current sample endorses culturally prevalent stereotypes about gender, race, and income might differ from the way a more general population endorses such stereotypes. These limitations are important to bear in mind when interpreting the results.

#### 4.6 Conclusions and implications

In summary, the main findings of the current study are that significant stereotype threat effects existed (1) on math performance on the basis of race and income-level, but not on the basis of gender and (2) on working memory function on the basis of gender, race, and income-level. In both cases, the income-based effects were the strongest. Further, there was evidence of multiple minority stereotype threat in that individuals with three stigmatized aspects of identity experienced significantly larger stereotype threat effects on math performance and working memory function than those with zero-, one-, or two-stigmatized aspects of identity.

These findings have particular relevance for researchers and educators. First, they highlight the power of income-based stereotype threat effects, a type of stereotype threat that is often overlooked by researchers and educators. Second, they highlight the complexity of identity and the number of aspects of identity that are sensitive to stigma. Considering this, stereotype threat researchers need to be sensitive to their participants' entire identity composition, not just the aspect of identity relevant to the type of stereotype threat being investigated in any given research project. Similarly, educators need to be mindful of the entire composition of their students' identities and aware of the confounding affects particular phrases and attitudes might have on the cognitive functioning and performance of the students in their classrooms.

**Acknowledgments** This research was supported by a Kaminsky Family Fund award and the Paul K. Richter and Evalyn E. Cook Richter Memorial Fund.

#### References

- Alloway, T. P. (2007). *Automated working memory assessment*. London: Harcourt Assessment.
- Baddedly, A. D., & Hitch, G. L. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.
- Beilock, S. L., Jellison, W. A., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2006). On the causal mechanism of stereotype threat: Can skills that don't rely heavily on working memory still be threatened? *Personality and Social Psychology Bulletin*, 32(8), 1059–1071. doi:[10.1177/0146167206288489](https://doi.org/10.1177/0146167206288489).
- Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology*, 136(2), 256–276. doi:[10.1037/0096-3445.136.2.256](https://doi.org/10.1037/0096-3445.136.2.256).
- Brooks-Gunn, J., Guo, G., & Furstenberg, F. (1993). Who drops out of and who continues beyond high school? A 20-year follow-up of black urban youth. *Journal of Research on Adolescence*, 3(3), 271–294. doi:[10.1207/s15327795jra0303\\_4](https://doi.org/10.1207/s15327795jra0303_4).
- Brown, C. S., & Leaper, C. (2010). Latina and European American girls' experiences with academic sexism and their self-concepts in mathematics and science during adolescence. *Sex Roles*, 63, 860–870. doi:[10.1007/s11199-010-9856-5](https://doi.org/10.1007/s11199-010-9856-5).

- Brown, R. P., & Day, E. A. (2006). The difference isn't black and white: stereotype threat and the race gap on Raven's advanced progressive matrices. *Journal of Applied Psychology, 91*(4), 979–984. doi:[10.1037/0021-9010.91.4.979](https://doi.org/10.1037/0021-9010.91.4.979).
- Bosker, R. J., Snijders, T. A. B., & Guldmond, H. (2003). PINT: Power analysis in two-level designs [software]. Available from <http://stat.gamma.rug.nl/snijders>.
- Cozzarelli, C., Wilkinson, A. V., & Tagler, M. J. (2001). Attitudes towards the poor and attributions for poverty. *Journal of Social Issues, 57*(2), 207–227. doi:[10.1037/0022-3514.69.5.797](https://doi.org/10.1037/0022-3514.69.5.797).
- Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*(6), 588–594. doi:[10.1177/0146167298246003](https://doi.org/10.1177/0146167298246003).
- Davies, P. G., Spencer, J. S., Quinn, D. M., & Gerardstein, 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. doi:[10.1177/014616702237644](https://doi.org/10.1177/014616702237644).
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5–18. doi:[10.1037/0022-3514.56.1.5](https://doi.org/10.1037/0022-3514.56.1.5).
- Forbes, C. E., & Schmader, T. (2010). Retraining attitudes and stereotype to affect motivation and cognitive capacity under stereotype threat. *Journal of Personality and Social Psychology, 99*(5), 740–754. doi:[10.1037/a0020971](https://doi.org/10.1037/a0020971).
- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28*(5), 659–670. doi:[10.1177/0146167202288010](https://doi.org/10.1177/0146167202288010).
- Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*, 17–28. doi:[10.1016/j.appdev.2007.10.004](https://doi.org/10.1016/j.appdev.2007.10.004).
- Gresky, D. M., Eyck, L. L. T., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles, 53*(9), 703–716. doi:[10.1007/s11199-005-7735-2](https://doi.org/10.1007/s11199-005-7735-2).
- Harrison, L. A., Stevens, C. M., Monty, A. N., & Coakley, C. A. (2006). The consequences of stereotype threat on the academic performance of White and non-White lower income college students. *Social Psychology of Education, 9*, 341–357. doi:[10.1007/s11218-005-5456-6](https://doi.org/10.1007/s11218-005-5456-6).
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155. doi:[10.1037/0033-2909.107.2.139](https://doi.org/10.1037/0033-2909.107.2.139).
- Jacobs, J. E., & Eccles, J. S. (1985). Gender differences in math ability: The impact of media reports on parents. *Educational Researcher, 14*(3), 20–25. doi:[10.2307/1174575](https://doi.org/10.2307/1174575).
- Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93*(4), 544–564. doi:[10.1037/0022-3514.93.4.544](https://doi.org/10.1037/0022-3514.93.4.544).
- Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology, 137*(4), 691–705. doi:[10.1037/a0013834](https://doi.org/10.1037/a0013834).
- 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. doi:[10.1023/A:1021003307511](https://doi.org/10.1023/A:1021003307511).
- Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin, 29*(3), 371–381. doi:[10.1177/0146167202250218](https://doi.org/10.1177/0146167202250218).
- 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*(6), 942–958. doi:[10.1037/0022-3514.80.6.942](https://doi.org/10.1037/0022-3514.80.6.942).
- Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80*(2), 713–732. doi:[10.1353/sof.2001.0102](https://doi.org/10.1353/sof.2001.0102).
- Lockhead, M. E., Thorpe, E., Brooks-Gunn, J., Casserly, P., & McAloon, A. (1985). *Understanding sex/ethnic related differences in mathematics, science and computer science for students in grades four to eight*. Princeton, NJ: Educational Testing Service.
- Lubienski, S. T. (2002). A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education, 71*(4), 269–287.
- McFarland, L. A., Lev-Arey, D. M., & Zeigert, J. C. (2003). An Examination of stereotype threat in a motivational context. *Human Performance, 16*(3), 181–205. doi:[10.1207/S15327043HUP1603\\_2](https://doi.org/10.1207/S15327043HUP1603_2).

- McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486–493. doi:10.1016/j.appdev.2006.06.003.
- Muller, C. (1998). Gender differences in parental involvement and adolescents' mathematics achievement. *Sociology of Education, 71*(4), 336–356. doi:10.2307/2673174.
- Nguyen, H. H. D., O'Neal, A., & Ryan, A. M. (2003). Relating test-taking attitudes and skills and stereotype threat effects to the racial gap in cognitive ability test performance. *Human Performance, 16*, 261–294. doi:10.1207/S15327043HUP1603\_5.
- Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat effect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314–1334. doi:10.1037/a0012702.
- Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement, 19*(1), 21–49. doi:10.1080/09243450801936845.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance, 16*(3), 231–259. doi:10.1207/S15327043HUP1603\_4.
- Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles, 59*, 377–391. doi:10.1007/s11199-008-9424-4.
- Quayle, M. (2011). Situated identity performance: Understanding stereotype threat as a social identity phenomenon. Doctoral dissertation, University of KwaZulu-Natal, Pietermaritzburg, South Africa.
- Raudenbush, S. W. et al. (2011). Optimal design software for multi-level and longitudinal research (version 3.01) [software]. Available from [www.wtgrantfoundation.org](http://www.wtgrantfoundation.org).
- Rosenthal, H. E. S., & Crisp, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality and Social Psychology Bulletin, 32*, 501–511. doi:10.1177/0146167205281009.
- Rydell, R. J., Beilock, S. L., & McConnell, A. R. (2009). Multiple social identities and stereotype threat: Imbalance, accessibility, and working memory. *Journal of Personality and Social Psychology, 96*(5), 949–966. doi:10.1037/a0014846.
- Rydell, R. J., & Boucher, K. L. (2010). Capitalizing on multiple social identities to prevent stereotype threat: The moderating role of self-esteem. *Personality and Social Psychology Bulletin, 36*(2), 239–250. doi:10.1177/0146167209355062.
- 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. doi:10.1037/0003-066X.59.1.7.
- Sawyer, T. P. Jr, & Hollis-Sawyer, L. A. (2005). Predicting stereotype threat, test anxiety, and cognitive ability test performance: An examination of three models. *International Journal of Testing, 5*, 225–246. doi:10.1207/s15327574ijt0503\_3.
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194–201. doi:10.1023/A:1018854212358.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440–452. doi:10.1037/0022-3514.85.3.440.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336–356. doi:10.1037/0033-295X.115.2.336.
- Scott-Jones, D., & Clark, M. L. (1986). The school experience of black girls: The interaction of gender, race, and socioeconomic status. *The Phi Delta Kappan, 67*(7), 520–526.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80–83. doi:10.1111/1467-9280.00111.
- Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research, 20*, 418–432. doi:10.1007/s11211-007-0047-7.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4–28. doi:10.1006/jesp.1998.1373.
- Smith, J. I. (2004). Understanding the process of stereotype threat: A review of meditational variables and new performance goal directions. *Educational Psychology Review, 16*(3), 177–206. doi:10.1023/B:EDPR.0000034020.20317.89.

- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. doi:10.1037/0022-3514.69.5.797.
- Steele, J., James, J. B., & Barnett, R. C. (2002). Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychology of Women Quarterly*, 26, 46–50. doi:10.1111/1471-6402.00042.
- Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, 16(1), 93–102. doi:10.1037/a0026617.
- Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, 34(4), 665–693. doi:10.1111/j.1559-1816.2004.tb02564.x.
- United States Department of Education, Institute for Education Sciences. (2009). *Percentage of students at or above selected mathematics proficiency levels, by age, sex, and race/ethnicity: Selected years, 1978 through 2008* [Data file]. Retrieved from [http://nces.ed.gov/programs/digest/d10/tables/dt10\\_141.asp](http://nces.ed.gov/programs/digest/d10/tables/dt10_141.asp).
- United States Census (2010). Quickfacts. Retrieved from <http://www.census.gov/>.
- Walsh, M., Hickey, C., & Duffy, J. (1999). Influence of item content and stereotype situation on gender differences in mathematical problem solving. *Sex Roles*, 41, 219–240. doi:10.1023/A:1018854212358.
- Wax, A. L. (2009). Stereotype threat: A case of overclaim syndrome? University of Pennsylvania Law School Public Law Research Paper No. 08–14.
- Wicherts, J. M. (2005). Stereotype threat research and the assumptions underlying analysis of covariance. *American Psychologist*, 60(3), 267–269. doi:10.1037/0003-066X.60.3.267.
- Willingham, W. W., & Cole, N. C. (1997). *Gender and Fair Assessment*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Yang, Y. (2003). Dimensions of socioeconomic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, 47(1), 21–41. doi:10.1111/j.1530-2415.2009.01184.x.

## Author Biographies

**Michele Tine** is an Assistant Professor in the Education Department at Dartmouth College. Her research program focuses on the ways socioeconomic status impacts malleable cognitive processes associated with academic achievement.

**Rebecca Gotlieb** is a 2012 graduate of Dartmouth College and is interested in integrating psychology, neuroscience and education to understand how social, emotional, and moral phenomena affect and are effected by learning and performance.