

Can Rewards Obviate Stereotype Threat Effects on Mental Rotation Tasks?

Amanda Kanoy, Sheila Brownlow, Tiffany F. Sowers
Department of Psychology, Catawba College, Salisbury, USA
Email: sbrownlo@catawba.edu

Received March 20th, 2012; revised April 21st, 2012; accepted May 24th, 2012

We examined whether sex-linked performance differences in Mental Rotation (MR) were obviated by rewards for performing the tasks. MR is typically seen as the domain of men, and therefore women completing the MR tasks likely worked under conditions of stereotype threat, which meant that their performance could vary according to situational variables. Men and women ($n = 33$ each) performed rotations and provided several self-reflective reports on their performances and background information about their experiences. Half of the participants (within sex) were rewarded for their participation with a gift card. Women's MR performance was lower than men's when no reward was given, but equaled it when they were rewarded. The finding was not a function of skill and self-reported effort, and emerged even when a stringent scoring technique was employed. The results suggest that rewards, even if they are not large, may nullify stereotype threat effects on women's MR.

Keywords: Stereotype Threat; Mental Rotation; Sex Differences

Can Rewards Obviate Stereotype Threat Effects on Rotation Tasks?

"In the end, it is impossible not to become what others believe you are" Gabriel Garcia Marquez in *Memories of My Melancholy Whores*, 2005.

Garcia Marquez's words illustrate the basic idea of *stereotype threat*: the fear people of clearly-defined groups have that they may confirm negative stereotypes about their group (Steele, 1997). Such fears often lead people to perform well below their level of competence (Steele & Aronson, 1995). If people claim group membership, understand the stereotype of the group, and are worried about what others will think of them personally (or their group as a whole), they can perform below their abilities on a variety of cognitive tasks (Shapiro, 2011; Shapiro & Neuberger, 2007).

There are many forms of stereotype threat that negatively impact people (Shapiro & Neuberger, 2007), but group-reputation threat is one that acutely influences individual behavior. This threat is the fear of confirming the negative group stereotype in the minds of the others outside the group (Schmader, Johns & Forbes, 2008). The concern goes beyond affect states and actually hinders cognitive performance, typically by increasing arousal (O'Brien & Crandall, 2003), evaluation concern (Brodish & Devine, 2009), self-handicapping behaviors (Keller, 2002), and by introducing intrusive thoughts about "letting your group down" (Cadinu, Maass, Frigerio, Impagliazzo, & Latiniotti, 2003; McIntyre, Paulson, & Lord, 2003).

Such thoughts may be at the root of the underperformance of Black men in comparison to their White peers on standardized tests (Aronson et al., 1998; Croizet & Claire, 1998), or for White men in similar circumstances when compared to Asian men (who may be seen as gifted in mathematics; Smith & White, 2002). While the earliest studies of how stereotype threat exerts its influence focused on the dimensions of race and

ethnicity, threat can also hinder performances of people based on other dimensions, including academic interest (Seibt & Förster, 2004), spatial relations (Brownlow, Valentine, & Owusu, 2008), social sensitivity (Koenig & Eagly, 2005), socioeconomic status (Croizet & Claire, 1998), and athleticism (Stone, Lynch, Sjomeling, & Darley, 1999).

Sex-linked stereotype threat effects have been well-documented in the domain of mathematics performance, an academic area where most believe—quite erroneously—that men are superior to women (Else-Quest, Hyde, & Linn, 2010). When experimental manipulations make clear that mathematics ability is to be tested, women do poorly in relation to men (Carr & Steele, 2009; Schmader, 2002; Spencer, Steele, & Quinn, 1999). The threat-induced decrease in math performance and other related cognitive tasks can be lessened or eliminated simply by reminding women of the positive, achievement-oriented aspects of their sex role (Schmader, Johns, & Barquissau, 2004), by highlighting another part of a social identity that is *not* deficient in math (Gresky, Ten Eyck, Lord, & McIntyre, 2005; Schmader et al., 2008), by presenting "peer testimonials" about the ease of the task (Brownlow, Janas, Blake, Rebadow, & Mellon, 2011), by claiming a test is being used simply to gather baseline information (Gonzalez, Blanton, & Williams, 2002), through the presentation of a high-achieving role model (Lesko & Corpus, 2006), and even by noting that women make better students and research subjects than do men (McIntyre et al., 2003). Thus, nullifying a stereotype threat and preventing underperformance is possible through a variety of means.

Stereotype threat affects self-efficacy, which in turn may influence actual task ability. If people perceive that they may fail at a given task, they may avoid the task (Spencer et al., 1999; Steele & Aronson, 1995). If they do engage and perform poorly, lack of efficacy is confirmed, perpetuating the idea of doubt about ability in the future (Schmader et al., 2004) and ultima-

tely leading to a lack of interest in that area (Keller & Dauheimer, 2003). Worse, stereotype threat can undermine actual ability, by preventing target persons from encoding and learning necessary information to start (R. Rydell, M. Rydell, & Boucher, 2010; Taylor & Walton, 2011).

Areas where women have little confidence and are underrepresented (such as STEM fields including science, engineering, and mathematics; see Shapiro & Williams, 2012) often employ Mental Rotation. Mental rotation (MR) is the transformation of three-dimensional blocks or objects in the head. Men outperform women by being quicker and more accurate at rotations (Bodner & Guay, 1997; Newcombe, 2007; Voyer, Voyer, & Bryden, 1995). Sex-linked differences in MR can be reduced or eliminated via practice, emphasis on accuracy over speed, or by shifting the focus on the tasks from rotations per se to generalized cognitive abilities (Alington, Leaf, & Monaghan, 2001; Sharps, Price, & Williams, 1994; Scali, Brownlow, & Hicks, 2000). Practice on the task (Kass, Ahlers, & Dugger, 1998), on spatial games (Cherney, 1998), and classes in mathematics and physical sciences are linked to better MR performance (Brownlow, McPherson, & Acks, 2003), as is athletic activity that employs spatial behavior (Ozel, Larue, & Molinaro, 2004; but only for men, Balentine & Brownlow, 2006).

In sum, stereotype threat provides one possible explanation for women's lack of MR ability in contrast to men. Although stereotype threat can be nullified in many ways, focusing on changing the cognitions or attributions for performance may be key to changing subsequent performance. Rewards may change not only behavior, but may also shift the attribution for behavior from internal to external (Freedman, Cunningham, & Krismer, 1992; Greenberg, Pyszczynski, & Paisley, 1984). Moreover, rewards for research participation imply that the task is difficult, unpleasant, and tedious, thus making attributions easier to externalize (Freedman et al., 1992), and perhaps alleviating concern that individual performance will reflect on an entire group. Thus, the purpose of this experiment was to examine how rewards for participation would influence performance on MR. We hypothesized that women's performance would improve if they were unconcerned or less concerned about stereotype threat, and therefore predicted that reward would improve women's performance in this domain.

Method

Participants

A total of 66 college students ($n = 33$ men; $n = 33$ women), aged 17 - 22, participated for course credit in psychology or sociology courses. Participants were assigned randomly within sex to complete rotation tasks for research credit only, or for credit *and* a gift card reward, resulting in a 2×2 (Sex \times Reward) between-participant design. Because all participants earned research credit (a requirement), the gift card to the College bookstore—given at the start of the experiment and part of the recruitment into the experiment—served as the reward.

Dependent Measures—MR Performance

The men and women completed the Purdue Visualization of Rotations Test (PVRT; Bodner & Guay, 1997), a test composed of 20 multi-dimensional block rotations. In this test, each block is paired with an identical shape that has been rotated along two dimensions (such as tilted forward and turned right), and under

that is another shape by itself. The test then shows five different rotated options for the unpaired shape, with only one correct rotation that matches the rotation pattern of the original paired shapes. The participants could score from 0 to 20 depending on the number of correct responses (raw score). The adjusted score (the raw score minus the number of incorrect responses) was recorded to accommodate guessing (Goldstein, Haldane, & Mitchell 1990), and ranged from -20 to 20. Time on task in s was also recorded.

Dependent Measures—Self Reports of Skill, Efficacy, Effort, and Handicapping

Self-efficacy and performance expectations may positively influence performance under stereotype threat (Smith; 2006), but lack of efficacy and tendency to self-handicap may increase stereotype-threat based underperformance (Stone, 2002). Thus, participants completed several measures of their skill, efficacy, efforts, and the judgments of the task and their beliefs about it using separate bipolar scales with the endpoints of 1 to 7, each bounded by opposite-meaning endpoints. One such question was how hard participants tried, bounded by 1 *I didn't try very hard* and 7 *I tried very hard*. The specific questions are described below in the section titled "data reduction."

Dependent Measures—Background

Background measures were taken because certain activities give people practice with MR tasks (Cadinu et al., 2003; Voyer & Isaacs, 1993). These activities included the number of sciences courses, particularly organic chemistry (Bodner & Guay, 1997), mathematics, dance, organized sports, and art classes. Self-reports about abilities in these areas were assessed on 7-point scales (endpoints labeled 1 *not good at all* to 7 *very good*). Participants reported the number of hours they played video/interactive games (0 - 2, 3 - 6, 7 - 10, 11 - 14, 15+).

Procedure

After obtaining consent, we told participants that "you are about to complete several problem solving tasks that involve rotating multi-dimensional blocks." These instructions were also provided on the cover page of the dependent measures booklet. Those in the reward group were told, "for your time and efforts, we are giving you a gift card to the College Bookstore; it's yours to keep for participation." The gift card was then given to the participant (or not), sample rotations were provided, questions answered, and the participants were left alone in the cubicle with the instructions to ring the bell once they started the first MR, and to ring the bell again when they finished (timed, in s , using a stopwatch). The PVRT booklet was removed and the men and women were given a packet of background assessment and self-efficacy scales. Reports of self-efficacy, and self-handicapping were taken in two orders to reduce order effects. Debriefing occurred at a later time when all the data had been collected.

Results

Data Analysis

Each major performance measure was entered separately into 2×2 (Sex \times Reward) ANOVAs. Then, self-report measures of

performance and background were subjected to separate factor analyses; resultant factors related to performance were then employed as covariates in analyses to examine whether efficacy and/or background mitigated the joint influence of sex and reward on performance.

Effect of Sex and Incentives on MR Performances

To examine how rewards influenced MR of women and men, the raw scores, adjusted scores, and the time on task (in sec) were separately entered into 2 × 2 (Sex × Incentive) ANOVAs. The means and standard deviations from these analyses can be seen in **Table 1**.

There was no main effect of reward on raw score, $F(1, 62) < 1.00$, $MSE = 12.01$, ns , nor a main effect for sex, $F(1, 62) = 2.58$, $p = .11$; however the Sex × Reward interaction was significant, $F(1, 62) = 4.19$, $p < .05$, $\eta_p^2 = .06$. Post hoc Scheffé tests comparing the means of men and women within reward condition demonstrated that women who received no reward ($M = 11.47$, $SD = 3.09$) did significantly poorer than men who did not receive a reward ($M = 14.59$, $SD = 3.20$). However, women who had the reward ($M = 13.06$, $SD = 3.44$) performed on par with men who were rewarded ($M = 12.69$, $SD = 4.09$).

Adjusted scores showed the same pattern, with no main effect of incentive, $F(1, 62) < 1.00$, $MSE = 48.32$, ns , and sex, $F(1, 62) = 2.99$, $p = .09$. However, the Sex × Incentive interaction produced a marginally significant effect, $F(1, 62) = 3.66$, $p = .06$, $\eta_p^2 = .06$. The adjusted scores had more variability than did the raw scores, thus the effect was only marginal, but the means followed the same pattern as the raw score means. Women who were rewarded ($M = 5.94$, $SD = 6.78$) scored on par with men who were rewarded ($M = 5.62$, $SD = 8.33$). In contrast, when stereotype threat was in the air and no reward was offered, men ($M = 9.18$, $SD = 6.40$) outperformed women ($M = 2.94$, $SD = 6.16$). There were no main effects of sex or incentive, nor was there an interaction between sex and incentive, for time-to-complete the task, all $F_s(1, 62) \leq 1.24$, $MSE = 53739.22$, all $ps \geq .24$

Data Reduction

Reduction of the data was necessary because multiple measures of the same constructs may have been taken. A factor analysis with varimax rotation was calculated on judgments of task liking, efficacy, self-handicapping, and perception of performance. The analysis produced six factors that accounted for 70.04% of the variance. The factors and their loadings are reported in **Table 2**. These factors were named *Skill* (including ability, enjoyment, positive performance evaluation, lack of frustration, understanding of the difficulty of the task, and perception that the task was not “tricky”), *Self-Handicapping* (affirmative judgments of recent life pressure, school stress, and feeling rushed during the day of the experiment), *Effort* (belief the test is valid, amount of effort put forth), *Evaluation Apprehension* (nervousness, concern for evaluation), *Task (Dis)Liking* (lack of enjoyment of task), and *Reward Pressure* (pressure due to the presence of reward).

A second factor analysis with varimax rotation was performed on measures of self-reported background experiences in the task. The factor analysis was performed to reduce and combine overlapping measures of sports experience, perceived math and science skill, and artistic abilities. The analysis produced

Table 1. Means, SD, and Fs for MR performance measures as a function of sex and reward.

	Men	Women	Total
Raw			
Card	12.69 (4.09)	13.06 (3.44)	12.88 (3.72)
No Card	14.59 _a (3.20)	11.47 _b (3.09)	13.03 (3.48)
Total	13.67 (3.73)	12.24 (3.31)	12.95 (3.57)
	<i>F</i> Sex (S)	<i>F</i> Reward (R)	<i>F</i> S × R
	2.58	<1.00	4.19*
Adjusted			
Card	5.62 (8.33)	5.94 (6.78)	5.78 (7.47)
No Card	9.18 (6.40)	2.94 (6.16)	6.06 (6.95)
Total	7.45 (7.51)	4.39 (6.55)	5.92 (7.16)
	<i>F</i> Sex (S)	<i>F</i> Reward (R)	<i>F</i> S × R
	2.99	<1.00	3.66
Time			
Card	636.50 (277.89)	538.69 (153.82)	587.59 (226.46)
No Card	573.00 (266.02)	609.41 (250.16)	591.21 (235.48)
Total	603.79 (250.56)	575.12 (208.97)	589.45 (229.38)
	<i>F</i> Sex (S)	<i>F</i> Reward (R)	<i>F</i> S × R
	<1.00	<1.00	<1.00

Note: $N = 66$. *F*-values marked with * $p < .05$. Means marked with different subscripts within rows differ, $p < .05$. Raw score varies from 20 to 5; adjusted scores from -10 to 20.

three background factors that accounted for 59.49% of the variance in the data. Specific loadings are shown in **Table 3**. *Sport Skill* (self-reported sports ability, length of time playing non-College sports), *Math-Science Skill* (self-reported math skill and self-reported science skill), and *Dance Background* (length of dance training, avoidance of video games) were the factors produced.

Relationship of Self-Reports and Background Measures to MR Performance

Factor means, after reverse scoring as needed, were calculated. These means were then correlated with the performance measures (MR time, raw score, and adjusted score). The relationships among self-reported efficacy/enjoyment factors, academic and sports background, and performance can be seen in **Table 4**.

Table 2.
Results of factor analyses on self-reports efficacy measures.

Factor and Variance	Measures (Loadings)
Skill (24.29%)	Ability (.79)
	Test Trickiness (-.76)
	Enjoyment (.81)
	Performance Evaluation (.81)
	Frustration (-.76)
	Task Difficulty (.76)
Self-Handicapping (12.95%)	Recent Life Pressure (.73)
	Recent School Stress (.78)
	Feeling Rushed During Day of Experiment (.73)
	Belief the Test is Valid (.69)
Effort (9.87%)	Effort (.82)
Evaluation Apprehension (9.37%)	Nervousness (.58)
	Evaluation Concern (.85)
Task (Dis)Liking (6.82%)	Lack of Task Enjoyment (.91)
Reward Pressure (6.74%)	Reward Pressure (.90)

Note: $N = 66$. Factor Analysis is the result of Principal Component Analyses with varimax rotation.

Table 3.
Results of factor analyses on background measures.

Factor and Variance	Measures (Loadings)
Sports Skill (24.14%)	Self-Reported Sports Ability (.85)
	Length of Playing Sports Before College (.88)
	Self-Reported Math Skill (.86)
Math-Science Skill (18.40%)	Self-Reported Science Skill (.70)
	Length of Dance Training (.78)
Dance Background (16.95%)	Hours/week Playing
	Video Games (-.72)

Note: $N = 66$. Factor Analysis is the result of Principal Component Analyses with varimax rotation

In sum, self-reported skill and effort positively related to raw score, and sports background was negatively related to raw score. A similar pattern appeared with adjusted score, with math/science skill also showing a positive correlation with that measure. Finally, only self-reported effort was related (positively) to time.

Influence of Self-Reports of Efficacy and Background as Mediators of Sex and Nullification Effects on MR Performance

Factors related to each performance measure were employed as covariates in 2×2 (Sex \times Reward) ANCOVAs on time, raw score, and adjusted score in order to examine whether any would change or eliminate the patterns described previously. For raw score, there were three covariates: sports skill, effort,

Table 4.
Correlations between self-reported efficacy, background factors, and mr performance.

Factor	MR Performance		
	Time	Raw Score	Adjusted Score
Self-reported skills	.02	.43***	.45***
Self-handicapping	-.05	.10	.10
Effort	.26*	.29*	.29*
Self-consciousness	.07	.19	.21
Lack of test enjoyment	-.12	.09	.10
Reward	-.03	.01	.02
Sports skill	-.16	-.29*	-.28*
Choreography/dance	-.06	-.10	-.09
Math/Science Ability	.16	.24	.26*

Note: $df = 64$, * $p < .05$, ** $p < .01$, *** $p < .001$.

and self-reported skill. Of these, self-reported skill was significant, $F(1, 58) = 10.04$, $MSE = 9.54$, $p = .002$, $\eta_p^2 = .15$, but the others were not. The ANCOVA results, in parallel with the ANOVA findings, showed no main effects of sex or reward, however, the interaction was again significant, $F(1, 58) = 4.29$, $p = .04$, $\eta_p^2 = .07$. Three covariates were significant for the analysis with adjusted scores, all $F_s(1, 57)$ and $MSE = 35.33$: sports skill ($F = 4.07$, $p = .048$, $\eta_p^2 = .07$), self-reported skill ($F = 5.63$, $p = .021$, $\eta_p^2 = .09$), and math/science ability ($F = 5.15$, $p = .027$, $\eta_p^2 = .08$). Despite the significant covariates, the pattern of results for adjusted scores remained as previous, with no main effects of sex or reward. However, there was a significant interaction, $F(1, 57) = 5.35$, $p = .024$, $\eta_p^2 = .09$. Finally, effort was a significant covariate to time to complete the task, $F(1, 61) = 5.53$, $MSE = 50082.92$, $p = .02$, $\eta_p^2 = .08$, but no main effects nor the interaction in the ANCOVA were significant.

Discussion

The results revealed that reward can obviate MR performance differences between men and women. The influence of the reward as a means to equalize women’s MR performance with men’s remained when various background skills and self-reported effort were held constant, and was not a function of self-handicapping, evaluation apprehension, incentive, task liking, math/science skill, reward pressure, or certain background experiences. As per previous research, men outperformed women when there was no reward and the stereotype threat was in the air (Newcombe, 2007; Scali & Brownlow, 2001; Sharps et al., 1994).

That women and men performed on par when rewarded shows that rewards worked differently for each, essentially increasing women’s performance and decreasing men’s, and suggests that rewards might nullify the deleterious effects of stereotype threat on women’s MR performance, assuming that women but not men were working under threat. Rewards may have affected women but not men because incentives for research participation connote that a task might be hard (Freedman et al., 1992), and perhaps women (who may have already thought the task was going to be unpleasant) were not further negatively influenced, but men (for whom the task should not

have been seen as onerous) did come to see the task that way. Like stereotype threat mediators, reward can positively and negatively affect performance. How an individual views the incentive can influence how well he or she performs, much like the way self-efficacy can affect performance. If a person is offered money, then performance expectations may increase (Ostrove, 1978). More importantly, research enticements may make attributions for performance easier to externalize and decrease internal attributions for behavior (Freedman et al., 1992), a situation which may have benefited women by reducing fear of low performance.

The results also suggest that it is possible that the reward nullified the stereotype threat for women by decreasing the intrusive thoughts that often occur for the targets of stereotype threat. That spatial tasks are generally thought of as “male” may have allowed the men to avoid the evaluation apprehension and intrusive thoughts that women may have experienced due to stereotype threat (Ostrove, 1978; Raty & Kasanen, 2007; Schmader, 2002). The women who received a reward may have been prevented from thinking about the task by being given the reward to them upfront—or perhaps women simply perceived that they needed to work harder to justify their gift. These women could make external attributions but did not do so because they experienced no evaluation apprehension (which decreases performance; Brodish & Devine, 2009). As a result, it is possible that no intrusive thoughts about performance occurred and the stereotype threat toward their sex was ignored, resulting no sex differences on the MR tasks when reward was given.

That reward can at least temporarily reduce sex-linked differences in MR has implications for understanding how women may come to change their evaluations of their personal abilities. If persons working under stereotype threat have doubt about their abilities on a task, those doubts can have a negative influence on self-perceived *future* abilities on related tasks (Schmader et al., 2004). Such self-doubt contributes to a lack of interest in pursuing a given field even if ability is present (Keller & Dauenheimer, 2003), because *others* continue to hold the stereotype (Cheryan, 2012). Reward may be a mechanism to keep a level of engagement that might otherwise be lost. Although paying people to complete cognitive and technical tasks in an institution is not a practical solution to the problems associated with stereotype threat, there are other forms of reward, including explicit encouragement, that may function the same way. Such explicit acknowledgements can have consequences for women’s participation in STEM fields where their presence is lacking (Shapiro, 2012). Whether rewards work to obviate group-related differences under stereotype threat on other cognitive tasks is still an open question.

REFERENCES

- Alington, D. E., Leaf, R. C., & Monaghan, J. R. (2001). Effects of stimulus color, pattern, and practice on sex differences in mental rotations task performance. *The Journal of Psychology, 126*, 539-553.
- Aronson, J., Lustina, M. L., Good, C., Keough, K., Steele, C. M., & Brown, J. (1998). When white men can’t do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29-46. doi:10.1006/jesp.1998.1371
- Balentine, C. B., & Brownlow, S. (2006). Does making salient task relevance to group affiliation decrease the performance of male athletes on spatial tasks? *Psi Chi Journal, 11*, 37-44.
- Bodner, G. M., & Guay, R. B. (1997). The purdue visualization of rotations test. *The Chemical Educator, 2*, 1-18. doi:10.1007/s00897970138a
- Brodish, A. B., & Devine, P. G. (2009). The role of performance-avoidance goals and worry in mediating the relationship between stereotype threat and performance. *Journal of Experimental Social Psychology, 45*, 180-185. doi:10.1016/j.jesp.2008.08.005
- Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology, 76*, 244-257. doi:10.1037/0022-3514.76.2.246
- Brownlow, S., Janas, A. J., Blake, K. A., Rebadow, K. T., & Mellon, L. M. (2011). Getting by with a little help from my friends: Mental rotation ability after tacit peer encouragement. *Psychology, 2*, 383-370. doi:10.4236/psych.2011.24057
- Brownlow, S., McPheron, T. K., & Acks, C. N. (2003). Science background and spatial abilities in men and women. *Journal of Science Education and Technology, 12*, 371-380. doi:10.1023/B:JOST.0000006297.90536.7c
- Brownlow, S., Valentine, S. E., & Owusu, A. (2008). Women athletes’ mental rotation under stereotypic threat. *Perceptual and Motor Skills, 107*, 307-316. doi:10.2466/pms.107.1.307-316
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology, 33*, 267-285. doi:10.1002/ejsp.145
- Carr, P. B., & Steele, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology, 45*, 853-859. doi:10.1016/j.jesp.2009.03.003
- Cherney, I. D. (2008). Mom, let me play more computer games: They improve my mental rotation skills. *Sex Roles, 59*, 776-786. doi:10.1007/s11199-008-9498-z
- Cheryan, S. (2012). Understanding the paradox in math-related fields: Why do some gender gaps remain while others do not? *Sex Roles, 66*, 184-190. doi:10.1007/s11199-011-0060-z
- Croizet, J., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*, 588-594. doi:10.1177/0146167298246003
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103-127. doi:10.1037/a0018053
- Freedman, J. L., Cunningham, J. A., & Krismer, K. (1992). Inferred values and the reverse-incentive effect in induced compliance. *Journal of Personality and Social Psychology, 62*, 357-368. doi:10.1037/0022-3514.62.3.357
- Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual-spatial ability: The role of performance factors. *Memory and Cognition, 18*, 546-550. doi:10.3758/BF03198487
- Gonzalez, 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*, 656-670.
- Greenberg, J., Pyszczynski, T., & Paisley, C. (1984). Effect of extrinsic incentives on use of test anxiety as an anticipatory attributional defense: Playing it cool when the stakes are high. *Journal of Personality and Social Psychology, 47*, 1136-1145. doi:10.1037/0022-3514.47.5.1136
- Gresky, D. M., Ten Eyck, L. L., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women’s performance under mathematics stereotype threat. *Sex Roles, 53*, 703-715. doi:10.1007/s11199-005-7735-2
- Kass, S. J., Ahlers, R. H., & Dugger, M. (1998). Eliminating gender differences through practice in an applied visual spatial task. *Human Performance, 11*, 337-349. doi:10.1207/s15327043hup1104_3
- 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
- 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*,

- 371-381. doi:10.1177/0146167202250218
- Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles, 52*, 489-496. doi:10.1007/s11199-005-3714-x
- Lesko, A. C., & Corpus, J. H. (2006). Discounting the difficult: How high math-identified women respond to stereotype threat. *Sex Roles, 54*, 113-125. doi:10.1007/s11199-005-8873-2
- Márquez, G. G. (2005). *Memories of my melancholy whores*. New York: Vintage.
- McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*, 83-90. doi:10.1016/S0022-1031(02)00513-9
- Newcombe, N. (2007). Taking science seriously: Straight thinking about spatial sex differences. In S. J. Ceci, & W. M. Williams (Eds.), *Why Aren't More Women in Science* (pp. 69-78). Washington, DC: American Psychological Association.
- 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.
- Ostrove, N. (1978). Expectations for success on effort-determined tasks as a function of incentive and performance feedback. *Journal of Personality and Social Psychology, 36*, 909-916. doi:10.1037/0022-3514.36.8.909
- Ozel, S., Larue, L., & Molinaro, C. (2004). Relationship between sport and spatial imagery: Comparison of three group of participants. *Journal of Psychology, 138*, 49-64. doi:10.3200/JRPL.138.1.49-64
- Raty, H., & Kasanen, K. (2007). Gendered views of ability in parents' perceptions of their children's academic competencies. *Sex Roles, 56*, 117-124.
- Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology, 99*, 863-896. doi:10.1037/a0021139
- Scali, R. M., & Brownlow, S. (2001). Influence of instructional manipulation and stereotype activation on sex differences in spatial task performance. *Psi Chi Journal, 6*, 3-13.
- Scali, R. M., Brownlow, S., & Hicks, J. (2000). Gender differences in spatial task performance as a function of speed or accuracy orientation. *Sex Roles, 43*, 359-376. doi:10.1023/A:1026699310308
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194-201. doi:10.1006/jesp.2001.1500
- Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles, 50*, 835-850. doi:10.1037/0033-295X.115.2.336
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype effects on performance. *Psychological Review, 115*, 336-356. doi:10.1037/0033-295X.115.2.336
- Seibt, B. F., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology, 87*, 38-56. doi:10.1037/0022-3514.87.1.38
- Shapiro, J. R. (2011). Different groups, different threats: A multi-threat approach to the experience of stereotype threats. *Personality and Social Psychology Bulletin, 37*, 464-480. doi:10.1177/0146167211398140
- Shapiro, J. R., & Neuberg, S. T. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review, 11*, 107-130. doi:10.1177/1088868306294790
- Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles, 66*, 175-183. doi:10.1007/s11199-011-0051-0
- Sharps, M. J., Price, J. R., & Williams, J. K. (1994). Spatial cognition and gender: Instructional and stimulus influences on mental image rotation performance. *Psychology of Women Quarterly, 18*, 413-425. doi:10.1111/j.1471-6402.1994.tb00464.x
- Smith, J. L. (2006). The interplay among stereotypes, performance-avoidance goals, and women's math performance expectations. *Sex Roles, 54*, 287-296. doi:10.1007/s11199-006-9345-z
- Smith, J. L., & White, P. H. (2002). An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles, 47*, 179-190. doi:10.1023/A:1021051223441
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1998). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-27. doi:10.1006/jesp.1998.1373
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613-629. doi:10.1037/0003-066X.52.6.613
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 68*, 797-811. doi:10.1037/0022-3514.69.5.797
- Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology, 77*, 1213-1227. doi:10.1037/0022-3514.77.6.1213
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects of Black and White athletic performance. *Journal of Personality and Social Psychology, 77*, 1213-1227. doi:10.1037/0022-3514.77.6.1213
- Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines actual learning. *Personality and Social Psychology Bulletin, 37*, 1055-1067. doi:10.1177/0146167211406506
- Voyer, D., & Isaacs, M. (1993). *Sex differences in mental rotation: Role of practice and experience*. Toronto: Annual Meeting of the Canadian Society for Brain, and Cognitive Science.
- Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250-270. doi:10.1037/0033-2909.117.2.250