Learning and Individual Differences 32 (2014) 273–277

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Implicit gender–math stereotype and women's susceptibility to stereotype threat and stereotype lift

Giulia Franceschini *, Silvia Galli, Francesca Chiesi, Caterina Primi

NEUROFARBA—Section of Psychology, University of Florence, Italy

ARTICLE INFO

Article history:
Received 20 March 2013
Received in revised form 28 November 2013
Accepted 23 March 2014

Keywords:
Implicit gender–math stereotype
Stereotype threat
Stereotype lift
Mathematics self-efficacy

ABSTRACT

This study explored the effects of implicit gender–math stereotyping on women's math self-efficacy and mathematics performance under stereotype threat and stereotype lift conditions. It was conducted with a sample of female undergraduate students enrolled in an introductory statistics course. Results showed that girls with implicit gender–math stereotype were sensitive to a stereotype threat-lift manipulation, whereas girls with weak implicit stereotype were not. Data suggest that implicit gender–math stereotyping acts as a critical variable in determining women's math self-efficacy and performance. These findings give some suggestions about the improvement of the teaching of math and related disciplines to female students.

1. Introduction

Although participation of girls and women in advanced mathematics studies, as well as in math-related activities and careers, has increased over the last decades (Hyde, Lindberg, Linn, Ellis, & Williams, 2008), the gender gap is still present. Cheryan (2012) argued that male and female gender roles can help in understanding why women choose not to enter math related careers. Indeed, math-related careers, especially in engineering and computer science, are stereotyped as masculine and are viewed as incongruent with female gender roles (e.g., Diekman, Brown, Johnston, & Clark, 2010) that traditionally include being socially skilled, helping others, having a family and raising children. Thus, among several explanations on which factors might contribute to gender differences in mathematics domain, the *Stereotype Threat* theory (Steele, 1997; Steele & Aronson, 1995) offers a perspective rooted in the power of social stereotype to influence thought and behavior.

The stereotype threat is one of the most heavily studied topics in social psychology over the past decade and refers to the concern that is experienced when one feels at risk of confirming, as self-characteristic, a negative stereotype about one's group (Steele & Aronson, 1995). The stereotype threat has been applied to the underperformance of African–Americans (e.g., on standardized tests: Steele & Aronson, 1995), white males (e.g., on athletic performances: Stone, 2002), Latinos (e.g., on difficult math tests: Gonzales, Blanton, & Williams, 2002), and a variety of other minority groups. In each case, the threat of confirming the stereotype undermines the performance of stigmatized individuals in the specific domain in which the stereotype applies.

Stereotypes concerning gender and mathematics ability propose that women have less mathematical aptitude than men. A growing body of research evidence indicates that these gender–mathematics stereotypes influence women's interest and performance in the mathematics domain (Davies, Spencer, Quinn, & Gerhardstein, 2002; Jacobs & Eccles, 1992; Quinn & Spencer, 2001; Sekaquaptewa & Thompson, 2003; Shih, Pittinsky, & Ambady, 1999; Spencer, Steele, & Quinn, 1999). Experimental studies suggest that gender differences in mathematics performance occur in environments evoking the gender stereotypes (Beilock, 2008), and that making participants' gender identity salient can be sufficient to create a threatening environment in which females' performance on demanding mathematics tasks drop to a lower-than-optimal level (Shih et al., 1999). This effect is especially impressive since it might apply in educational contexts. For example, Shapiro and Williams (2012), discussing how teachers' mathematics confidence might impact students' mathematics performance, illustrated the complexity of the mechanisms through which adults' mathematics attitudes can affect children. In the same way parental attitudes, and endorsement of gender stereotypes about mathematics, seem to be important for the development of girls' expectations of their own mathematics performance (Eccles, 2006). For example, Tomasetto, Alparone, and Cadinu (2011) found the moderating role of mother's gender stereotypes on girls' vulnerability to stereotype threat.

Along with the effects on performance, the stereotype threat theory (Aronson & Steele, 2005; Steele & Aronson, 1995) has been proposed to explain also gender differences in mathematics self-efficacy, defined as the self perception about one's capabilities to successfully perform a

* Corresponding author at: NEUROFARBA—Section of Psychology, Via di San Salvi 12—Padiglione 26, 50135 Firenze, Italy. Tel.: +39 0552055867.
E-mail address: franceschini.giu@gmail.com (G. Franceschini).

http://dx.doi.org/10.1016/j.lindif.2014.03.020
1041-6080/

mathematics task and obtain a satisfactory mathematics performance (Hackett & Betz, 1989; Pajares & Miller, 1994). Mathematics self efficacy – including students' belief about their own capabilities to successfully accomplish the subject in general, the specific tasks, and the ability to managing and monitoring their own learning – acts on performance in mathematics (Alenezi, 2008; Ayotola & Adedeji, 2009; Skaalvik & Skaalvik, 2004). While several evidences that stereotype threat interferes with mathematics performance have been provided, little research has explored the changes that this threat induces in mathematics self-efficacy. To the best of our knowledge, only Good, Aronson, and Harder (2008) have found that when the stereotype threat is nullified (i.e., when women were told that past results did not confirm the stereotype), women tend to perceive themselves more efficacious in solving mathematics problems than those ones in the stereotype threat condition (i.e., women who were told that past results confirmed the stereotype).

In the same way, despite the large body of literature examining the effects of stereotype threat on women's mathematics performance, little research has been done to examine the effects of *stereotype lift*. Traditionally the stereotype lift referred to members of groups favored by societal positive stereotype (e.g., men or Asian people in mathematics) but, alternatively, it refers to the boosting of performance in a given domain that occurs when an outgroup is negatively stereotyped or labeled (Walton & Cohen, 2003). Indeed, when cognizant of the weaknesses of an outgroup, an individual may approach the task more highly endorsing the assumption that he or she can succeed (Chalabaev, Stone, Sarrazin, & Croizet, 2008). Concerning gender–mathematics stereotype, Johnson, Bernard-Brak, Saxon, and Johnson (2012) adopted this alternative definition and compared three different conditions: threat, lift, and no threat/lift. They found that women perform better under no threat/lift or stereotype lift conditions than under stereotype threat, confirming that there is a detrimental effect on performance when the outgroup's superiority is stressed. Inversely, they did not find any difference between stereotype lift and no threat/lift conditions, suggesting that information about the outgroup's inferiority did not produce a boosting effect on women' mathematics performance.

Finally, another factor that we have to take into account is implicit processing. First, under stereotype threat, targets do not reliably report concerns about the stereotype when questioned directly (Steele & Aronson, 1995), and even when people report their explicit concerns about being stereotyped, the data does not reliably mediate stereotype threat effects (Bosson, Haymovitz, & Pinel, 2004; Johns, Schmader, & Martens, 2005; Wheeler & Petty, 2001). Thus, it is possible that stereotype threat effects occur, at least some of the time, without conscious awareness (Kiefer & Sekaquaptewa, 2007a). Second, a very recent study (Galdi, Cadinu, & Tomasetto, 2013) showed that girls make automatic associations consistent with the gender–mathematics stereotype from childhood, and it has been reported that women who possess implicit gender–mathematics stereotypes have less explicit mathematics identification and lower reported performance on mathematics-related achievement tests (Kiefer & Sekaquaptewa, 2007b; Nosek, Banaji, & Greenwald, 2002). Thus, if stereotype threat occurs through implicit processing of stereotype-relevant information, individual differences in the *implicit* gender–mathematics stereotyping, or in non-conscious cognitive associations of men with mathematics, may influence women's susceptibility to the stereotype threat. Kiefer and Sekaquaptewa (2007a) reported that when stereotype threat was reduced by describing the math test as non-diagnostic, the less women possessed implicit gender–math stereotypes, the better they performed. These findings were explained referring to the fact that women who possess strong implicit gender–math stereotypes have these stereotypes chronically accessible, and therefore may activate them even in the absence of stereotypic cues within the test-taking environment.

Starting from these premises, the present study aimed at investigating women mathematics self-efficacy and performance when exposed to stereotype threat and stereotype lift taking into account the role of *implicit* gender–mathematics stereotypes. That is, we examined how implicit gender–mathematics stereotyping might influence susceptibility to stereotype threat and lift referring both to performance and self-efficacy. In details, Kiefer and Sekaquaptewa (2007a, 2007b) manipulated the threat presenting a mathematics test as diagnostic or non-diagnostic. In the present study, to better understand if implicit gender–math stereotyping acts as a critical variable in determining women's math performance, we manipulated the threat *explicitly* referring to men and women's mathematical abilities (e.g., Good et al., 2008; McIntyre, Paulson, & Lord, 2003; Spencer et al., 1999). If women possess strong implicit gender–math stereotypes and have these stereotypes chronically accessible, this fact might make them especially sensitive to information related stereotype. Thus, in contrast with Kiefer and Sekaquaptewa (2007a), we hypothesized that girls who do not hold strong implicit stereotypes may be less sensitive to stereotype-related information, i.e., their math performance should be basically unaffected by those information. Vice versa, girls with strong implicit gender–mathematics stereotypes would be diminished under stereotype threat when explicitly told that woman perform worst. Additionally, always referring *explicitly* to men and women's mathematical abilities, we aim at exploring the potential boosting effect of counter-stereotypical information. Johnson et al. (2012) did not found any effect of stereotype lift on performance. However, they did not take into account the implicit mathematics–gender stereotype. In the present study, we aimed at better investigating the effect of informing about the outgroup's inferiority on both self-efficacy and performance of girls with a different level of implicit-stereotyping. Again, we hypothesized that who do not hold strong implicit stereotypes may be less sensitive to stereotype-related information, whereas a lift effect might be observed in girls who hold strong implicit stereotypes since these information are especially salient for them. This hypothesis is on line with studies that have demonstrated that counter-stereotypical information (e.g., informing women that they generally perform more reliably than men in a large array of tasks, or referring to women who have succeeded in several domains) alleviated the effects of stereotype threat on women's mathematics test performance (McIntyre et al., 2003).

2. Method

2.1. Participants

In line with previous investigations on the role of implicit stereotype on performance (Kiefer & Sekaquaptewa, 2007a, 2007b), the present study was conducted with a sample of female students. Participants were 78 female undergraduate students (mean age = 19.9, *DS* = 2.81) enrolled in an introductory statistics course at the University of Florence. Participants came from the following high schools: scientific lycée (23%), classic lycée (22%), linguistic lycée (10%), a teacher training school (32%), and industrial school (13%).

2.2. Measures and procedure

In order to measure mathematics skills and cognitive ability, participants received the *Prerequisiti di Matematica per la Psicometria* [Mathematics Prerequisites for Psychometrics] (PMP; Galli, Chiesi, & Primi, 2011) and the *Advanced Progressive Matrices—Short Form* (APM-SF; Arthur & Day, 1994). The PMP consists of 30 problems, and it has a multiple-choice format (one correct out of four alternatives). A single composite, based on the sum of correct responses, was calculated. The APM-SF is composed by 12 matrices derived from the APM, and its reliability and validity as a short-form of the Raven's Progressive Matrices has been tested (Chiesi, Ciancaleoni, Galli, Morsanyi, & Primi, 2012).

Additionally, participants completed individually the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998) administered on computers. The IAT assessed implicit gender–mathematics stereotyping and it consisted of seven stages in which participants were presented with a series of word categorization tasks. In stage 1, participants classified words presented individually on the computer screen as belonging to one of two categories, Humanistic and Scientific (with the majority of science-related concepts related to mathematics domains). If the word presented on the middle of the screen (e.g., "Literature") belonged to the category shown on the right of the screen (e.g., "Humanistic"), the participant responded by pressing the 'I' key. If the word flashed in the middle of the screen (e.g., "Mathematics") belonged to the category shown on the left of the screen (e.g., "Scientific"), the participant responded by pressing the 'E' key. In stage 2, a second pair of concepts was introduced (Male and Female). The participants again categorized words shown in the middle of the screen (e.g., "Father," "Wife") according to whether they belonged to the category shown on the right (e.g., "Male") or the left (e.g., "Female"). In stages 3 and 4, the categories from the first two stages were superimposed. The 'I' key was used to respond to words that referred to "Humanistic" or "Female," while the 'E' key was used for words that referred to "Scientific" or "Male." In stage 5, participants categorized words from the first stage but used the opposite keys to respond ("Humanistic", key "E"; "Scientific" key "I"). Finally in stages 6 and 7, the categories were again superimposed, but the key assignments for one pair were reversed from their position in the third and fourth stages ("Humanistic" or "Male" key "I"; "Scientific" or "Female" key "E"). Stages 1, 2, and 5 consisted of practice trials, whereas stages 3, 4, 6, and 7 consisted of test trials. Individual scores of automatic associations were calculated by means of the D-algorithm designed for analyzing data with the IAT (Greenwald, Nosek, & Banaji, 2003). The D-algorithm compares the extent to which performance on the incompatible critical block (i.e., Female/Science and Male/Humanistic sharing the same response key) is impaired relative to performance on the compatible critical block (i.e. Female/Humanistic and Male/Science sharing the same response key). The scoring takes into account participants' individual response latencies, standard deviations of latencies and error rates in each of the test blocks. Higher positive scores reflect stronger Male/Scientific and Female/Humanistic automatic associations.

Mathematics self-efficacy and performance were measured using the *Solution of Mathematics Problems* (SMP) subscale of the *Mathematics Self-Efficacy Scale revised* (MSES-R; Kranzler & Pajares, 1997; Pajares & Miller, 1995). This subscale consist of 18 mathematics problems for which students were asked to rate their confidence in solving them, using a five-point scale ranging from 1 (*no confidence*) to 5 (*complete confidence*). Participants were then asked to solve the 18 problems of the SMP. A single composite, based on the sum of correct responses, was calculated. High scores corresponded to high competences in mathematics.

2.3. Design

Participants were administered the PMP, APM-SF, and IAT. After, they were randomly assigned to stereotype threat vs. stereotype lift conditions. Under the stereotype threat condition participants read and were told: "You're going to take a test designed to evaluate certain skills in mathematics required in this final phase of the course. This type of test was used previously, and comparing the results of males and females was showed that *women are less able than men*." The stereotype lift condition was identical except for the sentence "*women are more able than men*". After the manipulation, students were presented the SMP scale. At the end of the experiment, participants were fully debriefed and thanked for their participation.

Thus, the study took the form of a 2 × 2 design with implicit gender–mathematics stereotype (recoded as Weak *vs* Strong, as detailed below) and stereotype manipulation (Threat *vs* Lift) as factors. The dependent variables were mathematics self-efficacy and performance on the mathematics test.

3. Results

In line with the cut-off (.63) proposed by Greenwald et al. (2003) to identify a high level of gender stereotype, two groups were created on implicit gender–mathematics stereotype scores (Weak and Strong) by taking the median (.62) of the IAT scores of the sample as a cut-off. In order to check the homogeneity of the experimental groups, 2 × 2 ANOVAs were conducted on mathematics skills and cognitive ability with implicit gender–mathematics stereotype (Weak vs. Strong) and stereotype manipulation (Threat *vs* Lift) as between factors. No differences were found for both the dependent variables. Concerning mathematics skills, main and interaction effects were not significant (Stereotype: $F_{(1,61)} = 2.51$, $p = .12$, Stereotype manipulation: $F_{(1,61)} = 1.04$, $p = .31$, Stereotype × Stereotype manipulation: $F_{(1,61)} = 0.014$, $p = .91$) and, in particular, no differences were found between those who reported weak and strong implicit gender–mathematics stereotype in threat and lift conditions (Threat: $M_{Weak} = 23.13$, $SD_{Weak} = 4.16$, $M_{Strong} = 21.24$, $SD_{Strong} = 5.52$, $t(31) = 1.10$, $p = .28$; Lift: $M_{Weak} = 24.13$, $SD_{Weak} = 3.90$, $M_{Strong} = 22.50$, $SD_{Strong} = 4.03$, $t(30) = 1.16$, $p = .26$). In the same way, main and interaction effects were not significant for cognitive ability (Stereotype: $F_{(1,59)} = .40$, $p = .53$, Stereotype manipulation: $F_{(1,59)} = .29$, $p = .60$, Stereotype × Stereotype manipulation: $F_{(1,59)} = .40$, $p = .53$) and, in particular, no differences were found between those who reported weak and strong implicit gender–mathematics stereotype in threat and lift conditions (Threat: $M_{Weak} = 8.75$, $SD_{Weak} = 1.88$, $M_{Strong} = 9.41$, $SD_{Strong} = 1.87$, $t(31) = -1.01$, $p = .32$; Lift: $M_{Weak} = 8.80$, $SD_{Weak} = 2.34$, $M_{Strong} = 8.81$, $SD_{Strong} = 2.24$, $t(28) = 0.01$, $p = .99$).

In order to examine the effects of implicit gender–mathematics stereotype (Weak *vs.* Strong) and stereotype manipulation (Threat *vs.* Lift) on mathematics self-efficacy, a 2 × 2 ANOVA was conducted. There was not a significant main effect of implicit stereotype ($F_{(1,60)} = .05$, $p = .83$) and stereotype manipulation ($F_{(1,60)} = 3.66$, $p = .06$), whereas the interaction effect was significant ($F(1,60) = 3.87$, $p < .05$, $\eta^2 = .06$). In details, students who reported low levels of implicit gender–mathematics stereotype did not show a different level of mathematics self-efficacy in the two conditions, and students who have a high level of implicit gender–mathematics stereotype showed lower self-efficacy in the threat condition when compared to the lift condition (Table 1). Comparing weak and strong stereotyped participants, independent sample t-tests did not reach the statistical significance; however, due to the small sample size of the samples (ranging from 14 to 16 cases), we computed the effect sizes (Kline, 2004). We found a small effect size ($d = .40$) when comparing math self-efficacy in the stereotype threat condition, with strong implicit gender–mathematics group reporting lower self-efficacy, and a medium effect size ($d = .60$) when comparing math self-efficacy in the stereotype lift condition, with strong implicit gender–mathematics group reporting higher self-efficacy.

Then, we conducted a 2 × 2 ANOVA to examine the implicit stereotype and stereotype manipulation's effect on mathematics performance. Whereas there was not a significant main effect of implicit stereotype

Table 1
Independent sample *t*-tests and descriptives for the weak and the strong implicit stereotype groups in the stereotype threat and stereotype lift conditions.

	Manipulation						
	Stereotype threat		Stereotype lift				
	M	*SD*	*M*	*SD*	*t* (*df*)	*p*	*d*
Mathematics self-efficacy							
Weak	80.53	12.15	80.38	11.80	0.37 (*31*)	.97	.01
Strong	75.40	13.23	86.76	9.31	−2.78 (*29*)	<.05	−.99
Mathematics performance							
Weak	10.94	2.86	10.88	2.94	0.60 (*30*)	.95	.02
Strong	8.64	2.82	11.57	2.31	−3.01 (*26*)	<.01	−1.13

($F_{(1,56)} = 1.26$, $p = .27$), a main effect of manipulation was found ($F_{(1,56)} = 4.04$, $p < .05$, $\eta^2 = .07$), and also the interaction effect was significant ($F_{(1,56)} = 4.40$, $p < .04$, $\eta^2 = .07$). In details, whereas students with weak stereotype did not show a different level of mathematics performance from threat to lift condition, students with strong stereotype showed lower mathematics performance in the threat condition when compared to the lift condition (Table 1). Comparing weak and strong stereotyped participants, we found a strong effect size ($d = .81$) when comparing the math performance in the stereotype threat condition, with strong implicit gender–mathematics group performing worse, and a small effect size ($d = .26$) when comparing the math performance in the stereotype lift condition, with strong implicit gender–mathematics group performing better.

4. Discussion and conclusion

This research aimed to investigate the effects of stereotype threat and stereotype lift in the domain of mathematics. On one hand, and consistent with the stereotype threat theory (Steele, 1997), women can suffer the effects of the threat showing a decrease in mathematics performance when it is made salient that women are less competent than men. On the other hand, in stereotype lift condition (Chalabaev et al., 2008; Walton & Cohen, 2003) when cognizant of the weaknesses of men, a woman may approach the task more highly endorsing the assumption that she can succeed. Given that stereotype threat and lift might occur through implicit processing of stereotype-relevant information and that the *implicit* gender–mathematics stereotype may influence women's susceptibility to stereotype-related information, the present paper aimed at investigating the relationship between the implicit gender–mathematics stereotype and the susceptibility to the stereotype threat and lift. Specifically, we looked not only at performance, as usually done in literature, but also at mathematics self-efficacy.

Results indicated that only female students who hold a strong implicit gender mathematics stereotype were susceptible to stereotype manipulation; in particular, they showed lower mathematics self-efficacy levels in the threat condition than in the lift condition. Thus, it seems that implicit gender mathematics stereotype plays a role in both threat and lift situations. Additionally, we found the same results for mathematics performance. More specifically, girls with strong implicit gender mathematics stereotype reported lower mathematics performances in the threat condition and a boost of the mathematics performance in the lift condition. Also in this case, the implicit gender mathematics stereotype seems to be an important predictor of the threat and lift effects.

This research provides a contribution to the study of implicit gender stereotype and to the study of stereotype lift. Thus, the current findings suggest that a strong implicit gender–mathematics stereotype makes girls susceptible to stereotype manipulations. Indeed, only girls who have strong associative links between the concepts of men and mathematics domains were more susceptible to stereotype threat and stereotype lift. When the stereotype was salient, we observe the detrimental effects on mathematics performance and self-efficacy. When the stereotype is activated given counter-stereotyped information, a lift effect is observed probably because women feel reassured that, at list in that specific case, they are more capable in mathematics than their male counterpart. Additionally, women who did not hold the stereotype did not change their performance moving from the threat condition to the lift condition, suggesting that the vulnerability of women is strongly related to their implicit gender–mathematics stereotype. We believe that these results emerged because implicit stereotyping increases the salience of gender-related information acting as a critical variable in determining women's math self-efficacy and performance.

The present findings might help in improving the teaching of mathematics and the related disciplines (such as statistics, physics, etc.) to female students. Indeed, this work provided concrete guidance on how to implement the lift effects or to prevent the threat effects through simple counter-stereotypical information. Given that it is possible to increase the confidence and to promote the performance encouraging expectations of success and emphasizing the ingroup capabilities, educational figures have an important role in contrasting the transmission of gender–mathematics stereotype and they might contribute to reduce gender differences in this critical sector.

This work fitted within a large body of research about the phenomenon of stereotype threat in women outside the laboratory setting and it offers some evidences about the relationships among implicit gender–mathematics stereotype, stereotype threat and lift on mathematics self-efficacy and performance. Nonetheless, several points are left uninvestigated. Specifically, future researches should include, on one hand, a control group to check the effects in the absence of the threat and lift conditions and, on the other hand, to include male participants in order to better investigate the effects of threat and lift on mathematics self-efficacy and mathematics performance (Johnson et al., 2012). Additionally, it might be interesting to investigate the relationships among threat and lift and implicit gender–mathematics stereotype in students belonging to STEM (Science, Technology, Engineering, Mathematics) faculty, in order to investigate the phenomena in people who daily face mathematics.

References

Alenezi, D. (2008). *A study of learning mathematics related to some cognitive factors and to attitudes.* (Doctoral dissertation). University of Glasgow.

Aronson, J., & Steele, C. M. (2005). Stereotypes and the fragility of human competence motivation, and self concept. In C. Dweck, & E. Elliot (Eds.), *Handbook of competence and motivation*. New York: Guilford.

Arthur, W., & Day, D. (1994). Development of a short form for the Raven Advanced Progressive Matrices test. *Educational and Psychological Measurement*, *54*, 395–403.

Ayotola, A., & Adedeji, T. (2009). The relationship between mathematics self-efficacy and achievement in mathematics. *Procedia Social and Behavioral Sciences*, *1*, 953–957.

Beilock, S. L. (2008). Mathematics performance in stressful situations. *Current Directions in Psychological Science*, *17*, 339–343.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology*, *40*, 247–255.

Chalabaev, A., Stone, J., Sarrazin, P., & Croizet, J. (2008). Investigating physiological and self-reported mediators of stereotype lift effects on a motor task. *Basic & Applied Social Psychology*, *30*, 18–26.

Cheryan, S. (2012). Understanding the paradox in mathematics-related fields: Why do some gender gaps remain while others do not? *Sex Roles*, *66*, 184–190.

Chiesi, F., Ciancaleoni, M., Galli, S., Morsanyi, K., & Primi, C. (2012). Item response theory analysis and differential item functioning across age, gender and country of a short form of the advanced progressive matrices. *Learning and Individual Differences*, *22*, 390–396.

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

Diekman, A. B., Brown, E., Johnston, A., & Clark, E. (2010). Seeking congruity between goals and roles: A new look at why women opt out of STEM careers. *Psychological Science*, *21*, 1051–1057.

Eccles, J. S. (2006). Families, schools, and developing achievement-related motivations and engagement. In J. E. Grusec, & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 665–691). New York: Guilford.

Galdi, S., Cadinu, M., & Tomasetto, T. (2013). The roots of stereotype threat: When automatic associations disrupt girls' mathematics performance. *Child Development*, http://dx.doi.org/10.1111/cdev.12128.

Galli, S., Chiesi, F., & Primi, C. (2011). Measuring mathematics ability needed for "non-mathematics" majors: The construction of a scale applying IRT and differential item functioning across educational contexts. *Learning and Individual Differences*, *21*(4), 392–402.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, *28*, 659–670.

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level mathematics courses. *Journal of Applied Developmental Psychology*, *29*, 17–28.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition. The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Hackett, G., & Betz, N. E. (1989). An exploration of the Mathematics Self-Efficacy/Mathematics performance correspondence. *Journal for Research in Mathematics Education*, *20*, 261–273.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize mathematics performance. *Science*, *321*, 494–495.

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, *63*(6), 932–944.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: teaching stereotype threat as a means of improving women's mathematics performance. *Psychological Science*, *16*, 175–179.

Johnson, H. J., Bernard-Brak, L., Saxon, T. F., & Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education*, *80*(2), 137–149.

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit stereotypes and women's mathematics performance: How implicit gender–mathematics stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*(5), 825–832.

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes, gender identification, and mathematics-related outcomes. *Psychological Science*, *18*(1), 13–18.

Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research.* Washington, DC: American Psychological Association.

Kranzler, J., & Pajares, F. (1997). An exploratory factor analysis of the Mathematics Self-Efficacy Scale Revised (MSES-R). *Measurement and Evaluation in Counseling and Development*, *29*(4), 215–228.

McIntyre, R. B., Paulson, R., & Lord, C. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*, 83–90.

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

Pajares, F., & Miller, M. (1994). Role of self-efficacy beliefs and self-concept beliefs in mathematics problem-solving: A path analysis. *Journal of Educational Psychology*, *86*, 193–203.

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics outcomes: The need for specificity of assessment. *Journal of Counseling Psychology*, *42*, 190–198.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematics problem-solving strategies. *Journal of Social Issues*, *57*, 55–71.

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, *39*, 68–74.

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*(3–4), 175–183.

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

Skaalvik, S., & Skaalvik, E. M. (2004). Frames of reference for self-evaluation of ability in mathematics. *Psychological Reports*, *94*, 619–632.

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

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

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

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

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girl's mathematics performance under stereotype threat: The moderating role of mothers' gender stereotype. *Developmental Psychology*, *47*, 943–947.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*, 456–467.

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