'Women are Bad at Math, but I'm Not, am I?' Fragile Mathematical Self-concept Predicts Vulnerability to a Stereotype Threat Effect on Mathematical Performance

FRIEDERIKE X. R. GERSTENBERG[1]*, ROLAND IMHOFF[2] and MANFRED SCHMITT[1]

[1]*University of Koblenz-Landau, Germany*
[2]*University of Bonn, Germany*

Abstract: The present research reports the results of three studies showing that individuals with a fragile self-concept in the domain of performance are particularly vulnerable to stereotype threat effects. Specifically, women who explicitly described themselves as rather mathematical but whose implicit self-concepts contradicted these claims were vulnerable to stereotype threat effects on mathematical performance. This effect was robust across three studies, independent of the subtleness or content of the stereotype threat manipulation. Additionally, it was shown that the effect was mediated by anxious worrying (Study 3). Copyright © 2012 John Wiley & Sons, Ltd.

A growing body of research has shown that individuals who are members of stereotyped groups suffer from a performance deficit when a negative stereotype about their in-group becomes salient (e.g., Croizet & Claire, 1998; Steele & Aronson, 1995). According to the Stereotype Threat (ST) model (Steele & Aronson, 1995), individuals who perform a difficult task in an area in which their group is considered weak feel at risk of confirming the stereotype, and this psychological pressure leads them to underperform. Although there is little doubt about the pervasiveness of the ST phenomenon, not all individuals are equally susceptible to its debilitating effects (e.g., Pinel, 2002; Schmader, 2002; Steele, Spencer, & Aronson, 2002). In the present paper, it is argued and shown that ST may particularly affect those with a fragile self-concept in the domain of performance (i.e., an explicit identification that is not backed up by a consistent automatic association).

Stereotype threat has been identified as a pervasive phenomenon. Although initially observed among African Americans taking an intellectual test and thus facing the threat of confirming a negative societal stereotype about their group's intellectual abilities (Steele & Aronson, 1995), other social groups have shown similar declines in performance when negative stereotypes about their groups' abilities were made salient. These groups have included people from low socio-economic backgrounds (Croizet & Claire, 1998), women (Spencer, Steele, & Quinn, 1999), elementary and middle-school girls (Ambady, Shih, Kim, & Pittinsky, 2001) and even men, an arguably more socially dominant group (Aronson, Lustina, Good, Keough, Steele, & Brown, 1999). Beyond

intellectual performance, ST can also impair performance in intergroup contexts (Goff, Steele, & Davies, 2008) as well as in athletic contexts (Chalabaev, Stone, Sarrazin, & Croizet, 2008; Stone, 2002; Stone, Lynch, Sjomeling, & Darley, 1999). The processes that drive ST include anxiety (Blascovich, Spencer, Quinn, & Steele, 2001; Osborne, 2001; Spencer et al., 1999), intrusive thoughts (e.g., Steele & Aronson, 1995; Study 3), a shift toward caution (Croizet & Claire, 1998; Steele & Aronson, 1995), performance expectancy (Cadinu, Maass, Frigerio, Impagliazzo & Latinotti, 2003; Stangor, Carr, & Kiang, 1998) and disengagement and effort withdrawal (Cadinu, Maass, Lombardo, & Frigerio, 2006; Spencer et al., 1999).

Building on the well-established findings that there exists an effect of ST, more recent research has tried to identify individual vulnerabilities (for a review, see Maass & Cadinu, 2003). Vulnerability to ST has been shown to depend on the degree to which individuals identify with the stereotyped group (Schmader, 2002) or the performance domain (e.g., Marx & Roman, 2002), the individuals' consciousness of the stigma (Pinel, 2002), locus of control (Cadinu et al., 2003, 2006) and the importance that individuals attribute to the relevant performance domain (Steele et al., 2002). Concerning individuals' automatic associations, Kiefer and Sekaquaptewa (2007a, 2007b) provided evidence that a strong automatic association of *math* with *man* moderated the effects of ST on women's mathematical performance. Thus, women who held implicit stereotypes about women as less mathematically talented were more affected by an ST manipulation than women who did not (the authors found no effect for self–woman or self–math associations). The present research builds on this pioneering work regarding the role of implicit processes in ST by adding a consistency perspective based on earlier findings. Specifically, it is

*Correspondence to: Friederike Gerstenberg (neé Dislich), Department of Psychology, Technische Universität München, Lothstr. 17, 80335 München, Germany.
E-mail: dislich@me.com

argued that an explicit self-concept that is not backed up by a consistent implicit self-concept constitutes a fragile self-concept that exposes individuals to the danger of ST. We postulate that a woman who has an explicit self-concept as a mathematical rather than an artistic person but who automatically associates herself more strongly with arts than with math will be most affected by ST effects in the mathematical domain.

IMPLICIT AND EXPLICIT MATHEMATICAL SELF-CONCEPT

In line with previous definitions of the self-concept (e.g., Asendorpf, Banse, & Mücke, 2002; Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, 2002), we define the mathematical self-concept (MSC) as the total strength of connections between the concept of who one is and the concept of mathematics. These connections may take the form of either propositional reasoning including truth values (e.g., 'I am (not) good at math!') or rather automatic associations between the two concepts *self* and *math*. This differentiation is in line with current dual system models (e.g., Smith & DeCoster, 2000) and the assumption of two underlying modes of information processing (e.g., Strack & Deutsch, 2004). Such models postulate distinct effects of explicitly endorsed beliefs (e.g., about one's own mathematical abilities) on the one hand and rather automatically activated associations between the concept of the self and the concept of math on the other hand. Nosek, Banaji, and Greenwald (2002) used an Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) to assess the implicit MSC (implicit math/arts identity). It was the aim of the present paper to explore the interplay of such implicit and explicit representations of the MSC in ST effects.

Following from the assumption that largely independent processes influence implicit and explicit representations (e.g., Gawronski & Bodenhausen, 2006; Grumm, Nestler, & von Collani, 2009), it is conceivable that individuals have contradictory self-concepts (e.g., explicitly judge themselves as highly mathematical but show only a weak association between the self and mathematics on rather indirect measures such as an IAT). Such discrepancies between explicit and implicit self-representations have previously been shown to be related to lower well-being (Schröder-Abé, Rudolph, Wiesner, & Schütz, 2007; Study 2), enhanced processing of discrepancy-related information (Briñol, Petty & Wheeler, 2006) and reduced intellectual performance (Dislich, Imhoff, Banse, Altstötter-Gleich, Zinkernagel, & Schmitt, 2011). Discrepant explicit and implicit self-esteem has been associated with ego fragility. In particular, the combination of high explicit self-esteem with low implicit self-esteem (often labelled *fragile*; e.g., Schröder-Abé et al., 2007) was found to predict in-group bias and dissonance reduction (Jordan, Spencer, Zanna, Hoshino-Brown, & Correll, 2003; Studies 2 and 3), increases in conviction strength and estimated consensus following manipulations of uncertainty and failure (McGregor & Marigold, 2003; McGregor, Nail, Marigold, & Kang, 2005) and less positive interpretations of ambiguous

feedback (Schröder-Abé et al., 2007; Study 1). In addition, individuals with a combination of high explicit and low implicit self-esteem scored high on the personality dimension of narcissism (Jordan et al., 2003; for a similar phenomenon at the group level, see Golec de Zavala, Cichocka, Eidelson, & Jayawickreme, 2009), supporting the *mask model* of narcissism, which assumes that inflated explicit self-esteem in narcissists is accompanied by suppressed feelings of low self-esteem. Nevertheless, the results concerning the mask model of narcissism are rather mixed (Bosson, Lakey, Campbell, Zeigler-Hill, Jordan, & Kernis, 2008).

With these findings, we postulated that a fragile configuration of implicit and explicit self-concepts may serve as a key to understanding individual differences in the vulnerability to ST. For the present research, which was focused on the effect of ST on women's mathematical performance, we argue that when an explicit representation of the self as mathematically skilled is confronted by a negation (as implied by a negative stereotype about one's own group), individuals with a corresponding positive implicit association have a buffer against the negation (Cohen & Wills, 1985), whereas individuals with a negative implicit association have nothing to buffer this threatening information and may start to worry about or question how mathematically skilled they actually are. For individuals with an explicit representation of the self as not very mathematically gifted, the information that is potentially threatening to others does not contradict but rather confirms their self-views. In that sense, they have no reason to be negatively influenced by ST.

THE PRESENT RESEARCH

Previous research has consistently shown that exposure to the stereotype of women as less mathematically talented than men can impair (some) women's subsequent mathematical performance (e.g., Aronson et al., 1999; Cadinu, Maass, Rosabianca, & Kiesner, 2005; Croizet & Claire, 1998; Spencer, Steele, & Quinn, 1999). We conducted three studies on how this effect is moderated by configurations of an implicit mathematical self-concept (iMSC) and an explicit mathematical self-concept (eMSC). Despite the continuous nature of both implicit and explicit self-concepts, for reasons of comprehensibility, we rely on prototypical labels that are based on the terminology previously proposed in the domain of the explicit–implicit configuration of self-esteem (Schröder-Abé et al., 2007; Table 1). In this framework,

Table 1. The four possible combinations of implicit and explicit mathematical self-concept

		Implicit mathematical self-concept	
		High	Low
Explicit mathematical self-concept	High	Consistently high	Fragile
	Low	Damaged	Consistently low

individuals can have: (i) a consistently high MSC (high eMSC and high iMSC); (ii) a fragile MSC (high eMSC and low iMSC); (iii) a damaged MSC (low eMSC and high iMSC); or (iv) a consistently low MSC (low eMSC and low iMSC). Specifically, we hypothesized that fragile MSCs expose women to greater vulnerability to suffer from ST.

This hypothesis was tested in three independent studies with three different ST manipulations. In previous research, ST has been manipulated in many different ways, including the use of techniques that are extremely subtle (having participants indicate their stereotyped demographic; e.g., Croizet & Claire, 1998), moderately subtle (characterising the test as diagnostic of ability; e.g., Steele & Aronson, 1995) and not-so-subtle (having participants read a packet of articles claiming that a specific out-group is superior in the to-be-tested domain; e.g., Aronson et al., 1999). Furthermore, one can distinguish ST manipulations by the elements that are activated. In our first study, ST was manipulated in a not-so-subtle fashion in which the threat content was activated (it was explicitly stated that women perform worse than men on this task). In Study 2, ST was manipulated in a subtle fashion in which participants were asked to indicate their gender and in which gender identity was activated. In Study 3, ST was, as in Study 2, manipulated in a subtle fashion by presenting sexist versus gender-neutral jokes. Thus, in the last study, derogative prejudices toward women were activated.

STUDY 1

The main goal of Study 1 was to test the relation between the MSC and mathematical performance in a design with two conditions: one with a not-so-subtle manipulation of ST and a control condition.

Method

Participants
A total of 146 female students with a variety of majors at the University of Koblenz-Landau participated in exchange for monetary compensation (€5). Their average age was 21.5 years ($SD = 2.6$).

Constructs and measurement instruments
Implicit mathematical self-concept. First, participants completed an IAT on the MSC (iMSC; similar to the math/arts identity IAT employed by Nosek et al., 2002, and the math identification IAT employed by Kiefer & Sekaquaptewa, 2007a, 2007b). The contrasting category of arts (vs. math) was chosen on the basis of previously established validity concerning the relation between IAT scores and actual mathematical performance (Kiefer & Sekaquaptewa, 2007a, 2007b; Nosek et al., 2002). A second advantage of a classical IAT with a binary target concept (i.e., math vs. arts) lies in the current efforts to measure individual differences rather than experimental effects. In contrast to variants of the IAT that use single targets (e.g., Karpinski & Steinman, 2006), classical IATs have superior psychometric properties and most importantly, reliability (Teige-Mocigemba, Klauer, & Sherman, 2010). Thus, the current IAT contrasted math versus arts (e.g., algebra/equations vs. literature/classics) and self-related versus other-related stimuli. Target and attribute stimuli and task sequence for the MSC-IAT are presented in Table 2. The MSC-IAT was scored according to the improved scoring algorithm (D600) proposed by Greenwald, Nosek, and Banaji (2003), reflecting the mean reaction time difference between two critical blocks (initial combined task vs. reversed combined task) in individual effect size units. To estimate its internal consistency, two standardized difference scores were computed for the odd trials in both critical blocks (divided by

Table 2. Implicit Association Test of mathematical self-concept: task sequence and stimuli

| Block | No. of trials | Task | Response key assignment | | Items |
			Left key	Right key	
1	20	Attribute discrimination	Mathematics	Arts	Math, algebra, geometry, calculus, equations, computation, numbers, Newton, poetry, art, Shakespeare, dance, literature, novel, symphony, drama
2	20	Target discrimination	I	They	I, me, my, mine, self, they, you, yours, theirs
3	60	Initial combined task	Mathematics or me	Arts or they	Math, algebra, geometry, calculus, equations, computation, numbers, Newton, I, me, my, mine, self, poetry, art, Shakespeare, dance, literature, novel, symphony, drama, they, you, yours, theirs
4	20	Reversed target discrimination	They	I	They, you, yours, theirs, I, me, my, mine, self
5	60	Reversed combined task	Arts or I	Mathematics or they	Poetry, art, Shakespeare, dance, literature, novel, symphony, drama, I, me, my, mine, self, math, algebra, geometry, calculus, equations, computation, numbers, Newton, they, you, yours, theirs

Note: The original German stimuli can be obtained from the authors.

the pooled *SD* for those trials) and the even trials in both critical blocks (divided by the pooled *SD*). These two *d* scores were then used to estimate a Spearman–Brown corrected split-half correlation. Using alternating trials instead of the first versus the second block ensured that both halves were maximally comparable and that attribute and target trials were equally represented in each half (see also Schmukle & Egloff, 2006).

Explicit mathematical self-concept. The eMSC was assessed by the difference in self-reported identification with math (compared with arts; Nosek et al., 2002). The following three items assessed the explicit math/arts self-concept by measuring the subjective link between an individual's self and math/arts: (i) 'Do you consider yourself to be a more mathematical person?' (ii) 'Do you consider yourself to be a more artistic person?' and (iii) 'Do you consider yourself to be a more artistic or a more mathematical person?' The first two items were scored by preference ratings ranging from 0 (*not at all*) to 100 (*very much*), and the third item was scored from −100 (*artistic*) to 100 (*mathematical*). To calculate the score of the eMSC, we combined Item 3 with the difference score of Items 1 and 2 and calculated the arithmetical mean. High (low) scores indicate a strong (weak) explicit representation between the self and mathematics and thus a high (low) eMSC.

Mathematical performance. The math test consisted of 20 math problems used in a study by Keller (2002), some of which were based on the Third International Mathematics and Science Study (14 items; Martin & Kelly, 1996) and some of which were selected from the math portion of the Graduate Management Test (six items). All items used a multiple-choice format in which four or five response options were given. Item difficulty ranged between 0.35 and 0.70, indicating that the mathematical performance test was of moderate to moderately high difficulty.

Procedure
The study was conducted in the laboratory in group sessions of up to five individuals. Upon arrival, participants were seated at individual computer stations where they completed the IAT. Subsequent to the IAT, all participants reported their eMSC. Then, the math test was announced. At the beginning of the math test, participants were informed that we were interested in individual differences in math performance and that we wanted to test some relevant theoretical assumptions. Participants were told that they should try to solve as many test items as possible and that their data would be treated confidentially. After that, participants were randomly assigned to the ST and control conditions. ST was manipulated in the introduction of the math test. Half of the participants read that the test had been shown not to produce gender differences ('The following math test is a collection of questions that have been shown not to produce gender differences in the past. The average achievement of male participants was equal to the achievement of female participants'). The other half read that the test had been shown to produce gender differences ('The following math test is a collection of questions that have been shown to produce gender differences in the past. The average achievement of male participants was different from the achievement of female participants. Male participants typically outperform female participants on this task'). Finally, participants had 20 minutes to complete the math problems, and then, they were fully debriefed and thanked.

Results

The two experimental groups did not differ with regard to their iMSC and eMSC, $F < 1$. Internal consistencies and descriptive statistics for all variables, along with intercorrelations between the variables, are presented in Table 3. Reliabilities were satisfactory for all measures. Effect sizes for all studies are reported as f^2, that is, the proportion of systematic variance accounted for by the interaction relative to the unexplained variance in the criterion (Aiken & West, 1991). To determine whether eMSC and iMSC of individuals were related to the scores they achieved on the math test and how these relations differed by experimental condition, we

Table 3. Descriptive statistics and correlations for all variables in Studies 1–3

	Study 1			Study 2			Study 3			
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(4)
(1) iMSC (IAT, D measure)	–			–			–			
(2) eMSC	0.18*	–		0.19*	–		0.20*	–		
(3) Mathematical performance	0.35**	0.31**	–	0.35**	0.32**	–	0.43**	0.46**	–	
(4) Worry	–	–	–	–	–	–	−0.41**	0.28**	−0.31**	–
#	n.a.	3	20	n.a.	3	20	n.a.	3	20	10
α	.80	.85	.72	.81	.86	.73	.84	.88	.81	.88
M	−0.55	15.81	8.03	−0.53	14.02	7.84	−0.59	15.54	8.12	15.83
SD	0.28	3.43	3.12	0.25	3.12	3.34	0.41	2.09	3.09	2.56
Min	−1.41	6	3	−1.51	4	2	−1.66	12	2	11
Max	0.15	20	17	0.17	19	16	0.18	21	16	22

Note: N (Study 1) = 146; *N* (Study 2) = 147; *N* (Study 3) = 156.
#, number of items; α, internal consistency (Cronbach's alpha); iMSC, implicit mathematical self-concept; eMSC, explicit mathematical self-concept; IAT, Implicit Association Test.
*$p < .05$; **$p < .01$.

conducted a multiple moderated regression analysis. We first centered the IAT and eMSC scores and then multiplied these centered variables by each other, as well as with the ST condition (effect coded as -1 for the control condition and $+1$ for the ST condition), to create the cross-product vector representing the interaction between them (Aiken & West, 1991). Math test scores were then regressed on these variables. Concerning performance on the math problems, the regression analysis ($R^2 = .25$; Cohen's $f^2 = .03$) yielded a significant main effect of iMSC, $\beta = .27$, $t(125) = 3.09$, $p < .01$, and a significant main effect of eMSC, $\beta = .30$, $t(125) = 3.70$, $p < .01$. Thus, both eMSC and iMSC predicted unique parts of the variance in mathematical performance. More importantly for our reasoning, the predicted three-way interaction involving iMSC, eMSC and the ST condition was also significant, $\beta = .18$, $t(125) = 2.05$, $p < .05$. This indicated that the size or direction of the ST effect was contingent on the configuration of eMSC and iMSC.

To explore this interaction in more detail, simple slopes tests at values of one *SD* above and below the means of iMSC and eMSC, dependent on ST condition, were used (Cohen & Cohen, 1983). As predicted, a significant effect of ST condition was obtained only for participants with a fragile MSC (high eMSC and low iMSC), $\beta = -.42$, $t(125) = 2.39$, $p < .05$, showing a negative effect of ST on their performance behaviour. For all other combinations of eMSC and iMSC, no ST effect was found (Figure 1).

Discussion

To our knowledge, Study 1 reports the first attempt to explore the role of individuals' specific configurations of eMSC and iMSC when examining ST effects. The results are fully consistent with our theoretical reasoning. As predicted, fragile individuals (high eMSC and low iMSC) were
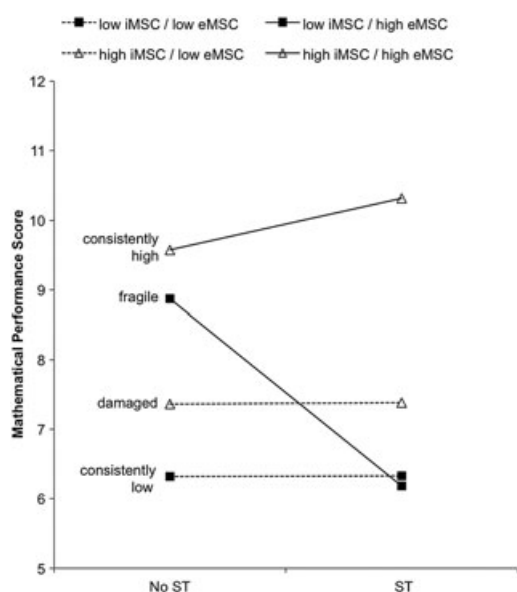

Figure 1. Predicted values for mathematical performance scores (Study 1) as a function of implicit mathematical self-concept and external mathematical self-concept after not-so-subtle versus no stereotype threat.

the only individuals to be affected by the ST condition. That is, fragile individuals in the ST condition performed significantly worse than fragile individuals in the control condition.

Furthermore, the analyses yielded significant main effects of both eMSC and iMSC on performance. This result may serve as an indication of the validity of both measures. Overall, participants who self-reported being mathematically skilled were more successful at solving mathematical problems than participants who reported being less mathematically skilled. Likewise, a stronger association between the self and the concept of math was associated with better math performance above and beyond the effect of the explicit claims. Thus, eMSC and iMSC additively predicted math performance.

In line with the assumption of a positive effect of a high iMSC, *damaged* MSC individuals (low eMSC and high iMSC) performed better than participants with a consistently low MSC in the control as well as in the ST condition, a pattern that supports the idea of a protective effect of high implicit self-associations. Thus, the results of Study 1 fully support our theoretical assumptions that individuals with a fragile MSC are affected by an ST manipulation but do not address the question of whether the effects can be replicated with a subtle ST manipulation in which only the concept of gender identification is activated. This question was addressed in Study 2.

STUDY 2

As discussed, the basic paradigm for demonstrating ST effects is one in which group members are made aware of a domain-relevant group stereotype and are then asked to take a test in that domain. In Study 1, the confrontation with the domain-relevant stereotype was rather blatant: Participants were explicitly told that their group (women) was expected to show suboptimal performance on the task at hand. Theoretically, this blatant manipulation could lead to demand effects because participants could be expected to see through the intentions of the researchers. Such demand effects could account for the results, under the assumption that participants with a fragile MSC were particularly inclined to show demand effects.

Although we had no reason to believe that participants were showing demand effects, we conducted a second study with a very subtle ST manipulation to bolster the generalizability of our findings. With the consistency of results from previous studies (e.g., Croizet & Claire, 1998), we hypothesized that the increase in the subtlety of the manipulation would not change the pattern of results from those predicted and found in Study 1.

Method

Participants

One hundred forty-seven women from different majors at the University of Koblenz-Landau participated in the study in exchange for monetary compensation. Their average age was 22.6 years ($SD = 2.3$).

Constructs and measurement instruments
All measures were the same as those used in Study 1.

Procedure
The procedure used in Study 2 was similar to the procedure used in Study 1. Only the ST manipulation was changed. In the subtle ST condition, participants had to indicate their gender directly before taking the math test (Croizet & Claire, 1998). Participants in the control condition did not indicate their gender.

Results

As in Study 1, the two experimental conditions did not differ with regard to their iMSC and eMSC, $F < 1$. Descriptive statistics and internal consistencies (Cronbach's alpha) for all measures as well as intercorrelations between all measures used in Study 2 are presented in Table 3. The internal consistencies of the measures were satisfactory.

Analyses were conducted in the same way as in Study 1. Concerning performance, the regression analysis ($R^2 = .28$; Cohen's $f^2 = .03$) replicated the significant main effects of iMSC, $\beta = .26$, $t(139) = 3.37$, $p < .01$, and eMSC, $\beta = .29$, $t(139) = 3.67$, $p < .01$, found in Study 1. Additionally, a significant two-way interaction between iMSC and eMSC, $\beta = .16$, $t(139) = 1.99$, $p < .05$, was qualified by the predicted significant three-way interaction involving iMSC, eMSC and the ST condition, $\beta = .18$, $t(139) = 2.30$, $p < .05$. Results again implied that the ST effect depended on the configuration of eMSC and iMSC.

More detailed analyses using simple slopes replicated the effects described in Study 1, showing a significant negative effect for individuals with a fragile MSC (high eMSC and low iMSC), $\beta = -.38$, $t(139) = 2.16$, $p < .05$ (Figure 2). No other significant ST effects were found. As in Study 1,
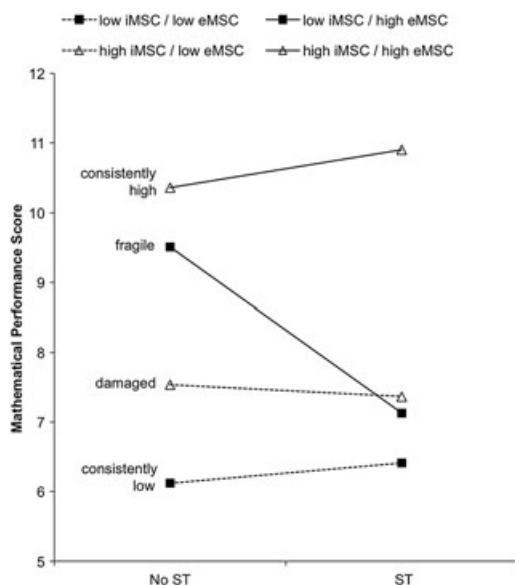

Figure 2. Predicted values for mathematical performance scores (Study 2) as a function of implicit mathematical self-concept and explicit mathematical self-concept after subtle versus no stereotype threat.

individuals with fragile MSC patterns performed worse after the ST manipulation than in the control condition, even when the ST manipulation was subtle.

Discussion

As in Study 1, the results of Study 2 are fully consistent with our theoretical reasoning and support our hypothesis. Only individuals with a fragile MSC (high eMSC and low iMSC) appeared to be affected by ST because they performed worse after a subtle ST manipulation compared with no manipulation. Also, in line with Study 1, no other configuration of eMSC and iMSC was affected by ST. Finally, Study 2 also replicated the main effects of eMSC and iMSC found in Study 1 and, thus, further corroborated the validity of the approach.

STUDY 3

Although the first two studies provided converging support for the hypothesis that ST manipulations are particularly harmful to the mathematical performance of individuals with a fragile MSC (high eMSC and low iMSC), in both studies, the ST condition was contrasted with a control condition in which participants either received no manipulation or were given no instructions before completing the math problems. Even if it is highly improbable that the observed effects can be attributed to this imbalance, we sought to replicate the findings and further extend them with a manipulation in which both experimental groups were given essentially the same task (rating jokes) that differed in their content: sexist (e.g., 'Why do women have four brain cells? One for each hot plate') versus nonsexist (e.g., 'Why is Windows not a virus? Viruses do not have major bugs'; Denzler, Förster, & Seibt, 2010). This manipulation has previously been shown to subtly activate stereotypes through jokes and as a result, to decrease speed on achievement tasks framed as intelligence tests for individuals affected by the stereotype (Denzler et al., 2010).

More importantly, we were interested in a potential mediating mechanism for the ST effect in individuals with a fragile MSC (high eMSC and low iMSC). We have argued that the worse performance of women with a fragile MSC may be because of their vulnerable self-concept, which leads them to worry more about performing poorly. This reasoning is in line with a plethora of findings that have established that ST effects are often mediated by an increase in worry or related constructs such as negative thinking (e.g., Beilock, Rydell, & McConnell, 2007; Brodish & Devine, 2009; Cadinu et al., 2005; Spencer et al., 1999). Schmader and Johns (2003) argue that ST interferes with performance by reducing the working memory capacity that individuals need to perform a task successfully. Working memory capacity can be thought of as a short-term memory system involved in the control, regulation and active maintenance of a limited amount of information with immediate relevance to the task at hand (Miyake & Shah, 1999). If the capacity of the working memory system to oversee task-relevant information is disrupted, performance may suffer. Steele and colleagues (2002) have suggested that ST is accompanied by 'concerns

about how one will be perceived, about one's ability, and thoughts about the stereotype' (p. 293). In support of this idea, women who were told that gender differences in math exist (i.e., stereotype threat) not only performed worse but also reported having more negative math-related thoughts compared with a control group (Beilock et al., 2007; Cadinu et al., 2005). Thus, situation-induced worry and verbal rumination may occupy central executive resources needed for integrating and monitoring the step-by-step processes of performance (Ashcraft & Kirk, 2001).

Worry refers to the cognitive elements of the anxiety experience, such as negative expectations and cognitive concerns about oneself, the situation at hand and potential consequences; this is in contrast to emotionality, which refers to one's perception of the physiological–affective elements of the anxiety experience (Liebert & Morris, 1967). In previous studies, threat of failure aroused worry but had no effect on emotionality (e.g., Morris & Liebert, 1973). Furthermore, the performance expectancies held by students as they enter test situations are highly related to their worry scores but unrelated to their emotionality scores (e.g., Morris & Liebert, 1970). In the context of the present research, we predicted that fragile individuals would have higher worry scores in the ST condition than in the control condition and in comparison to all other combinations of eMSC and iMSC. We further expected that worry would mediate the relation between MSC, ST condition and mathematical performance scores.

Method

Participants
One hundred fifty-six women with a variety of majors at the University of Koblenz-Landau were recruited to participate in this study in exchange for monetary compensation. Their average age was 24.5 years ($SD = 4.1$).

Constructs and measurement instruments
The measures of eMSC, iMSC and mathematical performance employed in Study 3 were identical to those used in Studies 1 and 2.

Worry. Worry was assessed using the worry subscale of the German Test Anxiety Inventory (Hodapp, 1991; English version: Hodapp, 1995; Hodapp & Benson, 1997). The subscale consists of six items such as 'I ask myself whether my performance will be good enough' or 'While taking a test, I often worry about how difficult it is.' Participants were asked to respond to each statement on a 4-point answer scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Higher scores indicate higher levels of worry.

Procedure
The procedure used in Study 3 was very similar to the procedure used in Studies 1 and 2, except that the ST manipulation was modified and the worry scale was included for Study 3. Participants, in sequence, completed the IAT, filled out the questionnaires on eMSC, rated jokes (sexist or nonsexist, depending on condition), completed the worry subscale of the German Test Anxiety Inventory and, finally, completed

the math test. In the ST-manipulation condition for Study 3, participants were asked to evaluate five sexist jokes, whereas participants in the control condition rated five neutral jokes. After completing all of the measures, participants were fully debriefed and thanked.

Results

As in Studies 1 and 2, the two experimental conditions did not differ with regard to their iMSC and eMSC, $F < 1$. The worry scores did not deviate from available norm data ($M = 15.99$, $SD = 4.49$), $t < 1$. Descriptive statistics and internal consistency statistics (Cronbach's alpha) for all measures used in Study 3 are reported in Table 3. The internal consistency levels of all of the measures were adequate. Table 3 also presents the intercorrelations of the measures. To determine whether eMSC, iMSC and ST condition were related to the mathematical performance scores and worry, multiple regression analyses were conducted.

Mathematical performance. Paralleling the analyses in Studies 1 and 2, mathematical performance scores were regressed on eMSC, iMSC, experimental condition and their interaction terms, $R^2 = .27$ (Cohen's $f^2 = .03$). Significant main effects were obtained for both iMSC, $\beta = .27$, $t(149) = 3.57$, $p < .01$, and eMSC, $\beta = .40$, $t(149) = 5.57$, $p < .01$. Supporting our hypothesis, these main effects were qualified by the three-way interaction involving experimental condition, eMSC and iMSC, $\beta = .18$, $t(149) = 2.12$, $p < .05$.

Simple slope analysis revealed a significant negative effect for individuals with a fragile MSC (high eMSC and low iMSC), $\beta = -.36$, $t(149) = 2.05$, $p < .05$ (Figure 3). Thus, the results of Studies 1 and 2 were replicated. Individuals with a fragile MSC performed worse after a subtle ST manipulation than they did after no ST manipulation. No other combinations of eMSC and iMSC were affected by ST.
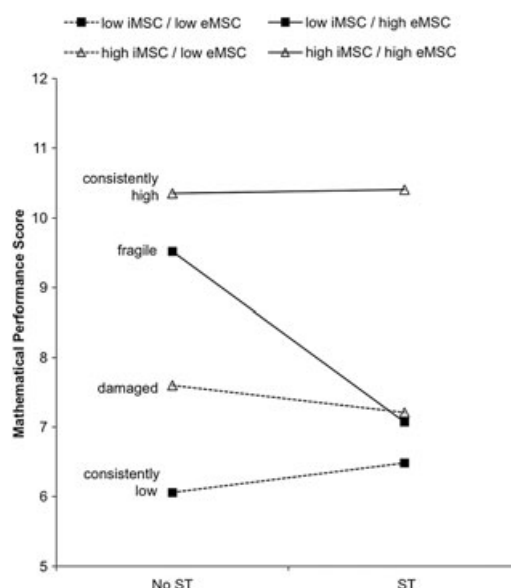
Figure 3. Predicted values for mathematical performance scores (Study 3) as a function of implicit mathematical self-concept and explicit mathematical self-concept after subtle versus no stereotype threat.

Worry. The hypothesized mediation of this effect through worry implied that participants with a fragile MSC would show not only the worst performance under ST but also the highest levels of worrying after the ST induction. To test this, we conducted a regression analysis parallel to the one reported above with the worry score as the criterion. The analysis, $R^2 = .29$ (Cohen's $f^2 = .04$), revealed significant main effects of iMSC, $\beta = -.38$, $t(149) = -5.28$, $p < .01$, and eMSC, $\beta = .36$, $t(149) = 5.26$, $p < .01$. These were qualified by a significant two-way interaction between ST condition and eMSC, $\beta = .16$, $t(149) = 2.26$, $p < .05$, as well as a significant two-way interaction between eMSC and iMSC, $\beta = -.23$, $t(149) = -2.86$, $p < .01$. Finally, these were further qualified by the predicted three-way interaction between ST condition, eMSC and iMSC, $\beta = -.19$, $t(149) = -2.11$, $p < .05$. The plotted slopes (Figure 4) show that despite the fact that participants with a fragile MSC already showed the highest worry scores without any ST induction, this worrying was further increased for the ST group (i.e., for those women who rated sexist jokes that implied that women are dull).

Simple slope analyses revealed a significant positive effect for individuals with a fragile MSC (high eMSC and low iMSC), $\beta = .39$, $t(149) = 2.18$, $p < .05$ (Figure 4). That is, individuals with a fragile MSC scored higher on the worry scale when taken after the ST manipulation than individuals in any other condition and with any other combination of eMSC and iMSC. No other ST effect was significant.

Moderated mediation analysis. To assess moderated mediation (Muller, Judd, & Yzerbyt, 2005; Preacher, Rucker, & Hayes, 2007), we examined the following four conditions: (i) significant effects of eMSC, iMSC and ST on math performance; (ii) significant effects between eMSC, iMSC and ST on worry; (iii) significant effects of worry on math performance controlling for the main and interaction effects of the MSC; and (iv) different conditional indirect effects of eMSC, iMSC and ST on math performance via worry, across low and high levels of eMSC, iMSC and ST. The last condition, which is the essence of moderated mediation, establishes whether the strength of the mediation via worry differed across low and high levels of eMSC, iMSC and ST.

The significant effects of eMSC and iMSC on math performance supported Condition 1 for moderated mediation (see aforementioned analyses). Likewise, the reported prediction of worry by the interactions of eMSC, iMSC and stereotype threat constituted Condition 2. Condition 3 was established through the significantly negative correlation between mathematical performance and worry ($r = -.31$) while controlling for the MSC. Hence, results based on the first three conditions indicate that eMSC, iMSC and ST condition moderate the mediation for worry on mathematical performance (Figure 5).

To further validate the findings of moderated mediation relations, we examined Condition 4, which requires the magnitude of the conditional indirect effect of mathematical performance via worry to be different for individuals with different MSC combinations. We used Preacher et al.'s (2007) statistical significance test, which applied Aroian's (1947) exact standard error for indirect effects, to compute a z statistic for the conditional indirect effect. We tested moderated mediation separately for each MSC combination dependent on experimental condition. Following Preacher et al.'s (2007) recommendation, we operationalized high and low levels of eMSC and iMSC as one SD above and below the mean score of the respective variable. The sampling distribution of the indirect effects was estimated through bootstrapping by resampling 10 000 units with replacement from the original sample of 156 units. For hypothesis testing, the null hypothesis of no indirect effect is rejected at the α level of significance if the value 'zero' lies outside the confidence interval. For the resulting nine configurations, significance tests were conducted on the null hypothesis that the conditional indirect effect equals zero. Results revealed that of all possible combinations of iMSC and eMSC, only for the fragile combination was there an indirect effect, indicated by the fact that zero was outside of the 95% confidence interval. Specifically, the region of significance had its lower bound at $-.9112$ and its upper bound at $-.0196$ with a mean conditional indirect effect of $\beta = -.46$, $p < .05$. No other conditional indirect effect was significant. Table 4 presents the beta weights, standard errors, lower and upper limits of the confidence intervals and the SDs of the conditional indirect effects for worry across consistently high, fragile, modest and consistently low MSC individuals.
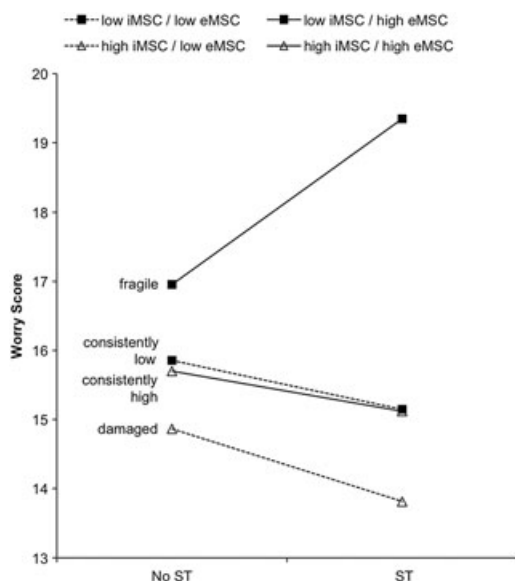

Figure 4. Predicted values for worry (Study 3) as a function of implicit mathematical self-concept and explicit mathematical self-concept after subtle versus no stereotype threat.

Discussion

Study 3 replicated the effect found in Studies 1 and 2 in a design that contrasted an ST manipulation with a very similar, nonthreatening manipulation. Again, the negative effect of
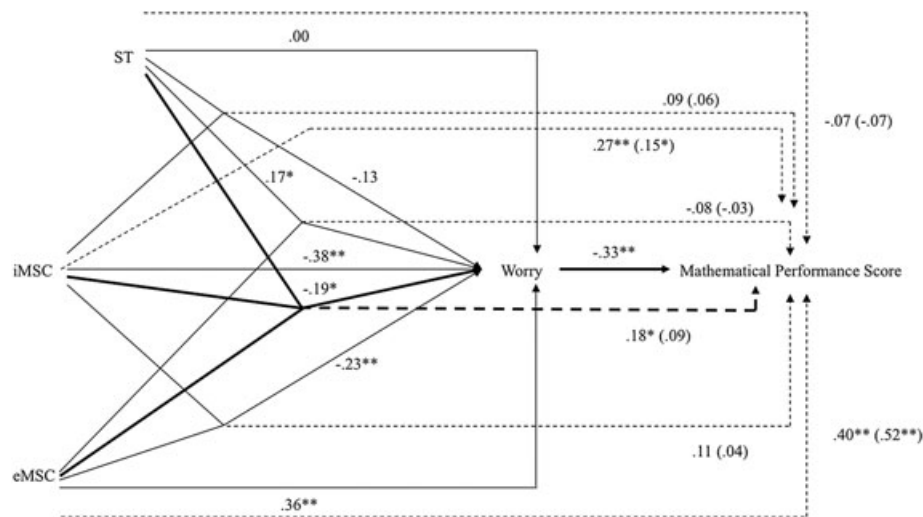
Figure 5. Test of the mediating role of worry for the effect of the mathematical self-concept (MSC) on the mathematical performance score in Study 3. Values are standardised regression weights. Dashed arrows depict main effects of implicit MSC; dotted arrows depict main effects of explicit MSC; bold arrows depict interaction effects between implicit MSC and explicit MSC; simple arrows depict main effects of worry. *$p < .05$, **$p < .01$.

Table 4. Bootstrapped conditional indirect effects of stereotype threat on mathematical performance through worry at specific values of the moderators (eMSC and iMSC; Study 3)

eMSC	iMSC	β	SE	LL CI	UL CI
−1 $SD = 13.46$	−1 $SD = -1.01$	.14	0.12	−0.1518	0.3310
−1 $SD = 13.46$	$M = -0.59$	.17	0.10	−0.0040	0.3572
−1 $SD = 13.46$	+1 $SD = -0.18$	.20	0.17	−0.0588	0.5527
$M = 15.54$	−1 $SD = -1.01$	−.16	0.12	−0.4337	0.0235
$M = 15.54$	$M = -0.59$	−.00	0.06	−0.1263	0.1213
$M = 15.54$	+1 $SD = -0.18$	.16	0.11	−0.0136	0.4020
+1 $SD = 17.63$	−1 $SD = -1.01$	−.46	0.23	−0.9112	−0.0196
+1 $SD = 17.63$	$M = -0.59$	−.18	0.12	−0.4254	0.0326
+1 $SD = 17.63$	+1 $SD = -0.18$	.11	0.13	−0.0995	0.4245

Note: N (Study 3) = 146.
eMSC, explicit mathematical self-concept; iMSC, implicit mathematical self-concept; β, beta weight; *SE*, standard error; LL CI, lower limit of the confidence interval; UL CI, upper limit of the confidence interval; *SD*, standard deviation.

ST on task performance was observed only for participants with a fragile MSC (high eMSC and low iMSC). Furthermore, we examined an integrated moderated mediation model to address the question of why fragile individuals were affected by stereotype threat. The findings of this analysis contributed to existing knowledge. Replicating the previous findings, evidence was obtained for a mediating effect of worry; participants with a fragile MSC had the highest scores on the worry scale. This greater degree of worry was found to be responsible for the depreciation in performance of these individuals.

As has been shown by other researchers (e.g., Cadinu et al., 2005; Steele & Aronson, 1995), performance-inhibiting thoughts or emotions, such as worry, appeared to mediate ST effects.

GENERAL DISCUSSION

Taken together, the findings of these three studies provide strong converging evidence that combinations of eMSC

and iMSC determine vulnerability to ST. Women with a congruently high MSC were not affected by either not-so-subtle or subtle ST manipulations, whereas women with a fragile MSC (high eMSC and low iMSC) showed the expected effect of poorer performance. It may be that ST effects are commonly underestimated because of a lack of differentiation; ST may affect only those individuals who have a fragile self-view in the performance domain at hand. The results of the present research suggest that ST operates via the interplay of individual and situational variables that are mediated by emotional factors such as anxious worrying. The studies presented here propose that an individual's MSC plays a key role in the person's experience of ST.

Judging from the influence that self-concept discrepancy had in our studies, it is of great interest to know more about how discrepant versions of the MSC develop. We have treated the eMSC and the iMSC as pre-existing individual differences in our current study. Although it is beyond the scope of the present paper to further explore the antecedents

of eMSC and iMSC, it is open to debate how a discrepant self-concept develops. In our research, we presume that individuals with a fragile MSC (high eMSC and low iMSC) tend to exaggerate their current and past performances. The repeated past experience of failure may have led individuals to have an automatically activated representation of being not quite as mathematically skilled as they explicitly claim to be. This underlying representation as less mathematically gifted than explicitly claimed exposes them to the threat of worry, ultimately impairing their actual performance. As our results suggest, the well-established effect of ST may in fact be driven by a performance decline that is displayed only by individuals with insecure (i.e., fragile) self-concepts. These compelling findings are particularly relevant for education policy and may help intervention programmes undermine the destructive effects of ST. However, fragile MSC individuals performed at high levels in the control conditions. Another explanation would be that women with a fragile combination of MSC are particularly concerned about confirming gender stereotypes regarding the mathematical abilities of women. It may be the case that these chronic concerns then lead to a low iMSC. Thus, future research should distinguish between identification with math and math abilities. For example, some women studying mathematics may identify strongly with mathematics but may question their mathematical abilities—particularly if they judge them in reference to other math students rather than in reference to the general population.

As a first limitation of the present studies, we have relied solely on female participants. In no study was a male control group included; thus, we are merely presuming that the observed performance decline from our manipulations was an ST effect. It might be argued that men with a fragile MSC could have reacted in the same way to our ST manipulations. However, we consider it unlikely that information about female underperformance, marking one's gender or reading sexist jokes would have the same debilitating effect on men with a fragile MSC. As our main interest was in the role of individual differences, we maximized statistical power of detecting those by focusing on the group of interest (i.e., women).

A second potential limitation is the use of a classical IAT and not of a single-category IAT (SC-IAT). Thus, the iMSC was inherently negatively confounded with the implicit artistic self-concept. It might be argued that SC-IAT (Karpinski & Steinman, 2006) could have circumvented this problem by constituting a purer measure of self–math associations. Nevertheless, we decided against an SC-IAT as previous evidence pointed to the (math-related) validity of classic math/arts IATs (e.g., Kiefer & Sekaquaptewa, 2007a, 2007b; Nosek et al., 2002) and under special consideration of the generally worse psychometric properties of SC-IATs. Particularly for research on individual differences, the reliability of the measure is of crucial importance. Thus, it may be seen as a major drawback that a high MSC is confounded with a low artistic self-concept. Furthermore, the mean index of implicit associations between self and mathematics was negative, and the maximum observed value of the index is just a little over

zero. Although this might lead some attentive readers to the speculation that for the absolute majority of our participants, the self was associated more with the arts than with math, we would like to add a caveat to help avoid a premature interpretation of this finding. To maximize the visibility of individual differences (not contaminated by compatibility order), we always presented the blocks in the same order with the self–math association first. As the second critical block is generally slower (Nosek, Greenwald, & Banaji, 2007), most likely because of task switching costs (Klauer & Mierke, 2005), a negative score does not indicate a stronger self-association with the arts than with mathematics per se. Although it is still conceivable that participants having (relatively) high iMSC in our study could be considered to have a neutral self-concept, we would like to stress that they still performed better than individuals with a lower mathematical self-concept, thus reflecting the validity of these individual differences.

It is also noteworthy that neither the IAT nor the explicit scales referred to individuals' mathematical abilities but rather to the degree to which math (vs. arts) is tied to themselves. Individuals might think of themselves as more mathematical than artistic because they think in mathematical terms, like to solve mathematical problems or because they are members of a family of mathematicians. These different alternatives do not imply that these individuals also think that they are good at math. Nevertheless, in our argument, we treated an MSC as at least partially related to (explicit and implicit) self-estimations of mathematical abilities. Although we have no strong evidence to test this, the main effects of both eMSC and iMSC across all three studies may be a first indication. Participants with strong (explicit as well as implicit) associations of the self with math performed better on a mathematical test. This indicates that it is highly plausible to assume the MSC to be (partially) related to mathematical ability.

In summary, our results corroborate previous findings regarding the vulnerability associated with fragile self-concepts (Briñol et al., 2006; Dislich et al., 2011; Schröder-Abé et al., 2007) in the domain of ST. Our results add not only to the important questions of who is particularly likely to suffer from ST but also to an underlying mechanism (distraction due to worry). These findings may help to more strongly focus interventions targeted at ST effects in the field. In light of limited resources, it seems useful to concentrate on individuals with fragile self-concepts in the relevant domain of performance rather than to apply interventions indiscriminantly to everyone.

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, *12*, 385–390. DOI: 10.1111/1467-9280.00371

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics*, *18*, 265–271. DOI: 10.1214/aoms/1177730442

Aronson, J., Lustina, M. J., Good, C. Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35*, 29–46. DOI: 10.1006/jesp.1998.1371

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380–393. DOI: 10.1037/0022-3514.83.2.380

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology. General*, *130*, 224–237. DOI: 10.1037/0096-3445.130.2.224

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spill over. *Journal of Experimental Psychology. General*, *136*, 256–276. DOI: 10.1037/0096-3445.136.2.256

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and blood pressure: The role of stereotype threat. *Psychological Science*, *12*, 225–229. DOI: 10.1111/1467-9280.00340

Bosson, J. K., Lakey, C. E., Campbell, W. K., Zeigler-Hill, V., Jordan, C. H., & Kernis, M. H. (2008). Untangling the links between narcissism and self-esteem: A theoretical and empirical review. *Social and Personality Psychology Compass 2/3*, 1415–1439. DOI: 10.1111/j.1751-9004.2008.00089.x

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, *91*, 154–170. DOI: 10.1037/0022-3514.91.1.154

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

Cadinu, M. R., 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

Cadinu, M. R., Maass, A., Lombardo, M., & Frigerio, S. (2006). Stereotype threat: The moderating role of Locus of Control beliefs. *European Journal of Social Psychology*, *36*, 183–197. DOI: 10.1002/ejsp.303

Cadinu, M. R., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*, 572–578. DOI: 10.1111/j.0956-7976.2005.01577.x

Chalabaev, A., Stone, J., Sarrazin, P., & Croizet, J.-C. (2008). Investigating physiological and self-reported mediators of stereotype lift effects on a motor task. *Basic and Applied Social Psychology*, *30*, 18–26. DOI: 10.1080/01973530701665256

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analyses for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*, 310–357. DOI: 10.1037/0033-2909.98.2.310

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*, 588–594. DOI: 10.1177/0146167298246003

Denzler, M., Förster, J., & Seibt, B. (2010). *Just a joke?* The effects of the dumb-blond stereotype on speed and accuracy. Manuscript submitted for publication.

Dislich, F. X. R., Imhoff, R., Banse, R., Altstötter-Gleich, C., Zinkernagel, A., & Schmitt, M. (2011). Discrepancies between implicit and explicit self-concepts of intelligence predict performance on tests of intelligence. *European Journal of Personality*. DOI: 10.1002/per.827

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731. DOI: 10.1037/0033-2909.132.5.692

Goff, P. A., Steele, C. M., & Davies, P. G. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology*, *94*. 91–107. DOI: 10.1037/0022-3514.94.1.91

Golec de Zavala, A., Cichocka, A., Eidelson, R., & Jayawickreme, N. (2009). Collective narcissism and its social consequences. *Journal of Personality and Social Psychology*, *97*, 1074–1096. DOI: 10.1037/a0016904

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25. DOI: 10.1037/0033-295X.109.1.3

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. DOI: 10.1037/0022-3514.74.6.1464

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. DOI: 10.1037/0022-3514.85.2.197

Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: The case of self-esteem. *Journal of Experimental Social Psychology*, *45*, 327–335. DOI: 10.1016/j.jesp.2008.10.006

Hodapp, V. (1991). Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten (The Test Anxiety Inventory TAI-G: An expanded and modified version with four components). *Zeitschrift für Pädagogische Psychologie*, *5*, 121–130.

Hodapp, V. (1995). The TAI-G: A multidimensional approach to the assessment of test anxiety. In C. Schwarzer, & M. Zeidner (Eds.), *Stress, anxiety, and coping in academic settings* (pp. 95–130). Tübingen, Germany: Francke.

Hodapp, V., & Benson, J. (1997). The multidimensionality of test anxiety: A test of different models. *Anxiety, Stress and Coping*, *10*, 219–244. DOI: 10.1080/10615809708249302

Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Brown, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology*, *85*, 969–978. DOI: 10.1037/0022-3514.85.5.969

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test (SC-IAT) as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*, 16–32. DOI: 10.1037/0022-3514.91.1.16

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

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit stereotypes and women's math performance: How implicit gender-math

stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*, 825–832. DOI: 10.1016/j.jesp.2006.08.004

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science 18*, 13–18. DOI: 10.1111/j.1467-9280.2007.01841.x

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the IAT effect. *Personality and Social Psychology Bulletin*, *31*, 208–217. DOI: 10.1177/0146167204271416

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*, 975–978.

Maass, A., & Cadinu, M. (2003). Stereotype Threat: When minority members underperform. *European Review of Social Psychology*, *14*, 243–275. DOI: 10.1080/10463280340000072

Martin, O. M., & Kelly, D. L. (Eds.). (1996). *Third international mathematics and science study: Technical report*. Chestnut Hill, MA: Boston College.

Marx, T., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin*, *28*, 1183–1193. DOI: 10.1177/01461672022812004

McGregor, I., & Marigold, D. C. (2003). Defensive Zeal and the uncertain self: What makes you so sure? *Journal of Personality and Social Psychology*, *85*, 838–852. DOI: 10.1037/0022-3514.85.5.838

McGregor, I., Nail, P. R., Marigold, D. C., & Kang, S.-J. (2005). Defensive pride and consensus: Strength in imaginary numbers. *Journal of Personality and Social Psychology*, *89*, 978–996. DOI: 10.1037/0022-3514.89.6.978

Miyake, A., & Shah, P. (1999). *Models of working memory. Mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.

Morris, L. W., & Liebert, R. M. (1970). Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance. *Journal of Consulting and Clinical Psychology*, *35*, 332–337. DOI: 10.1037/h0030132

Morris, L. W., & Liebert, R. M. (1973). Effects of negative feedback, threat of shock, and level of trait anxiety on the arousal of two components of anxiety. *Journal of Counseling Psychology*, *20*, 321–326. DOI: 10.1037/h0034768

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, *89*, 852–863. DOI: 10.1037/0022-3514.89.6.852

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59. DOI: 10.1037/0022-3514.83.1.44

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, *26*, 291–310. DOI: 10.1006/ceps.2000.1052

Pinel, E. C. (2002). Stigma consciousness in intergroup contexts: The power of conviction. *Journal of Experimental Social Psychology*, *38*, 178–185. DOI: 10.1006/jesp.2001.1498

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227. DOI: 10.1080/00273170701341316

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. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440–452. DOI: 10.1037/0022-3514.85.3.440

Schmukle, S. C., & Egloff, B. (2006). Assessing anxiety with extrinsic Simon tasks. *Experimental Psychology*, *92*, 149–160. DOI: 10.1027/1618-3169.53.2.149

Schröder-Abé, M., Rudolph, A., Wiesner, A., & Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology*, *42*, 174–183. DOI: 10.1080/00207590601068134

Smith, E. R., & DeCoster, J. (2000). Dual process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131. DOI: 10.1207/S15327957PSPR0402_01

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28. DOI: 10.1006/jesp.1998.1373

Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology*, *75*, 1191–1187. DOI: 10.1037/0022-3514.75.5.1191

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. DOI: 10.1037/0022-3514.69.5.797

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

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*, 1167–1678. DOI: 10.1177/014616702237648

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology*, *77*, 1213–1227. DOI: 10.1037/0022-3514.77.6.1213

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. DOI: 10.1207/s15327957pspr0803_1

Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). Practical guide to Implicit Association Tests and related tasks. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (p. 117–139). New York: Guilford Press.