Journal of Personality and Social Psychology
2006, Vol. 91, No. 2, 243–254

0022-3514/06/$12.00 DOI: 10.1037/0022-3514.91.2.243

Distinguishing Stereotype Threat From Priming Effects: On the Role of the Social Self and Threat-Based Concerns

David M. Marx and Diederik A. Stapel
University of Groningen

It has been argued that priming negative stereotypic traits is sufficient to cause stereotype threat. The present research challenges this assumption by highlighting the role of the social self and targets' concerns about confirming a negative group-based stereotype. Specifically, in 3 experiments the authors demonstrate that stereotype threat adversely affects the test performance and threat-based concerns of targets (but not nontargets) because only targets' social self is linked to the negative group stereotype. Trait priming, however, harms the test performance of both targets and nontargets but has no effect on their threat-based concerns because trait priming does not require such a link between the social self and the group stereotype. Moreover, the authors show that merely increasing the accessibility of the social self in nonthreatening situations leads to the underperformance of targets but has no meaningful effect on nontargets' test performance.

Keywords: stereotype threat, priming, threat-based concerns, social identity, performance

To date, considerable research has investigated the adverse effects of stereotype threat on targets' test performance (Steele, 1997; Steele, Spencer, & Aronson, 2002). This predicament, known as stereotype threat, refers to situations in which stereotyped targets underperform when a negative stereotype is relevant to their performance. For example, if women take a math test that they think is diagnostic of their mathematical ability, they typically have lower scores than men because the diagnosticity manipulation (see Steele & Aronson, 1995) makes the stereotype relevant to women's performance. When the same math test is described in nonthreatening terms, however, women and men perform similarly because the negative stereotype is irrelevant to women's performance. Stereotype threat is now well documented among a variety of groups and performance domains (e.g., Croizet & Claire, 1998; Gonzales, Blanton, & Williams, 2002; Inzlicht & Ben-Zeev, 2000; Marx & Roman, 2002; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995; Stone, Lynch, Sjomeling, & Darley, 1999). Given the ubiquity of stereotype threat effects, the next step is to clarify how stereotype threat differs from other processes that also lead to stereotype confirming behavior, namely stereotype priming effects. Because there is a growing body of research investigating priming and stereotype threat, it appears particularly important and timely to distinguish between these two accounts. Moreover, because the behavioral outcomes of priming and stereotype threat can look the same (poor performance), a better understanding of stereotype-based performance seems needed.

David M. Marx and Diederik A. Stapel, Department of Social and Organizational Psychology, University of Groningen, Groningen, the Netherlands.

This research was supported in part by "Pionier" Grant "Making Sense of Hot Cognition" from the Dutch National Science Foundation (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) and a research grant from the Heymans Institute of the University of Groningen awarded to Diederik A. Stapel. We thank Sei Jin Ko for her helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to David M. Marx, who is now at the Psychology Department, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-4611, or to Diederik A. Stapel, Department of Social and Organizational Psychology, University of Groningen, Grote Kruisstraat 2/1, Groningen 9712 TS, the Netherlands. E-mail: dmarx@sciences.sdsu.edu or d.a.stapel@rug.nl

Priming and Stereotype Threat

One of the primary aspects of stereotype threat is that a negative stereotype (e.g., "women are bad at math") must be activated to harm targets' test performance (e.g., women's poor math test performance). This has led some researchers to view the priming of negative stereotypic traits as a sufficient means to create a stereotype threat situation, because a stereotype threat situation is also a situation in which stereotype activation leads directly to stereotype-related behaviors (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004; Dijksterhuis & Bargh, 2001; Dijksterhuis & Corneille, 2004; Gladwell, 2005; Oswald & Harvey, 2000; Shih, Ambady, & Pittinsky, 1999; Wheeler, Jarvis, & Petty, 2001). As several experiments have shown, when participants are primed with labels or stereotypic traits, they behave in a stereotype-consistent manner, such as walking slower down a hallway when primed with the stereotype of an older person (Bargh, Chen, & Burrows, 1996) or performing less intelligently on general knowledge tests when thinking about soccer hooligans (Dijksterhuis & van Knippenberg, 1998).

In contrast to this priming perspective, which suggests that stereotype threat could be viewed as a general priming effect (i.e., stereotype activation → stereotype-confirming test performance), others have argued that stereotype threat is more than a mere priming effect. For instance, according to Marx, Brown, and Steele (1999), targets' poor test performance is due to the "situational pressure posed by the prospect of being seen or treated through the lens of a negative group stereotype" (p. 493). This situational pressure, in turn, increases targets' worry about being judged in

terms of the stereotype associated with their stereotyped identity if they underperform (e.g., Steele, 1997; Steele et al., 2002). Consider a situation in which women and men are taking a difficult math test. If the women perform poorly, this could be perceived as evidence for the stereotype that "women cannot do math." Consequently, in stereotype threat experiments, unlike in priming experiments, there is a clear relationship between the negative stereotype and targets' (but not nontargets') performance on the test. Even though there is a large amount of research exploring the effects of stereotype threat and priming on participants' performance, a direct empirical comparison between these two accounts of stereotype-based performance has not been made until now. Thus, at present, we do not know to what extent stereotype threat and stereotype priming effects are empirically similar or different (but see Wheeler and Petty, 2001, for a review of stereotype threat and priming effects on behavior).

In this article, we argue and demonstrate that stereotype threat is more than a general priming effect: Stereotype threat is a situational predicament that links one's performance to the concern about confirming a negative group stereotype (i.e., "I worry that my test performance may confirm the negative stereotype about my group"). In other words, although priming effects typically rely on perception-to-behavior effects that refer to relatively global, nonspecific links between activation and behavior (e.g., priming soccer hooligans causes aggressive behavior for anyone because this category label activates the relevant concept, aggressiveness), stereotype threat relies on perception-to-behavior effects that are less global and more specific. Stereotype priming can affect anyone, whereas stereotype threat, by definition, only occurs for those people who are targeted by the relevant stereotype. Of course, in both cases the stereotype needs to be activated, but it is how people view the stereotype that distinguishes between these two accounts of stereotype-based performance.

Thus, whereas stereotype knowledge (e.g., associating older people with the trait *slow* or associating hooligans with the trait *aggressive*) is sufficient for stereotype activation to result in general behavioral effects (e.g., walking slower, acting aggressively), such knowledge is not enough for stereotype threat effects to occur. For stereotype threat, both knowing and being are necessary (see Marx, Stapel, & Muller, 2005; Steele, 1997; Steele et al., 2002). For example, in stereotype threat situations, targets (but not nontargets) are affected because they know the group stereotype ("women are bad at math") and because they are members of the group that is targeted by the stereotype ("I am a woman"). Hence, self-relevance of a stereotype is one of the keys to stereotype threat effects. Priming self-relevant or other-relevant stereotypes can lead to stereotype-consistent behavior (e.g., Shih, Ambady, Richeson, Fujita, & Gray, 2002; Wheeler et al., 2001), but only when the stereotype is self-relevant may the primed behavior be related to the stereotype threat experience. Thus, on the surface, the differences between negative trait priming and threat appear somewhat indistinguishable: Both trait priming and threat can lead to the same outcome (poor performance). However, if one looks below the surface, the differences are clear. For trait priming effects, it does not matter who one is; for threat effects, it does. That is, if one is targeted by a negative stereotype, then one's concerns about confirming that stereotype are likely to be increased. Not only does stereotype threat induce stereotype-consistent behavior, it also induces worry and concern (Spencer et al., 1999; Steele et al., 2002). It is interesting that none of the research that has looked at stereotype threat from a priming perspective has included such a measure of worry or concern. Rather, the relevant research to date has focused mainly and primarily on test performance and not on participants' self-related, threat-based concerns (Ambady et al., 2004; Shih et al., 1999, 2002; Wheeler et al., 2001).

Threat-Based Concerns

Although at first glance the results from many priming experiments appear consistent with the notion that priming stereotypic traits may be sufficient for causing stereotype threat effects on targets' performance, we think that priming procedures are insufficient because they bypass one of the hallmarks of stereotype threat: the worry about confirming a negative stereotype associated with one's group (Major & O'Brien, 2005; Marx et al., 1999; Steele et al., 2002). It is also important to note that because much of the research exploring the impact of negative traits on performance has not included both targets and nontargets (see e.g., Ambady et al., 2004; Bargh et al., 1996; Dijksterhuis & van Knippenberg, 1998; Wheeler et al., 2001; but see Shih et al., 2002, for work on positive trait priming among targets and nontargets), it remains difficult to determine whether negative trait priming affects the performance of targets and nontargets in the same way and whether trait priming has any effect on participants' threat-based concerns (see also Wheeler & Petty, 2001).

We contend, therefore, that what is needed to assess this *knowing-and-being* aspect of stereotype threat are measures of participants' threat-based concerns, as this type of measure can show whether targets link their performance in the testing situation to their stereotyped identity. That no measure of participants' threat appraisals has been included in past priming research only helps to fuel the misperception that negative trait priming can elicit stereotype threat. By including such a measure, it may be possible to show that despite similar performance outcomes the primes have very different effects on participants' threat-based concerns (high concern = stereotype threat, low or no concern = priming).

The Role of the Social Self

This knowing-and-being logic of stereotype threat also suggests that stereotype threat situations are especially likely to increase the accessibility of the social identity that is most relevant to the stereotype. Recent research by Marx et al. (2005) provides an illustration of this point. These researchers showed that female participants' gender group (relative to other social groups, such as friends and family) is more accessible in a math stereotype threat situation than in a nonthreat situation. This finding shows that stereotype threat situations increase the accessibility of the relevant social self (gender).[1] In the present research, we test whether

[1] Some may view this reasoning as similar to research on identity bifurcation by Pronin, Steele, and Ross (2004). On the contrary, we see the Pronin et al. research and the present research as being rather distinct. Pronin et al. focused on how the situation (either a stereotype threat or neutral one) can lead to differences in targets' identification with certain characteristics that are stereotype relevant, whereas we focus on how one's social identity in combination with the situational cues can trigger stereotype threat (see also Marx et al., 2005).

the reverse is also true: Will increased accessibility of a generalized social self change a nonthreatening testing situation into a threatening one if targets are given a relevant situational cue (seeing math problems)? Our reasoning behind this notion is that increasing accessibility of a generalized social self ("we-ness") may thus enhance targets' threat-based concerns—even in nonthreatening testing situations—because the relevant cues likewise increase the accessibility of the group stereotype associated with the situation (Major & O'Brien, 2005; Marx et al., 1999, 2005; Steele et al., 2002).

Although we are not the first to discuss the role of the social self in stereotype-based performance effects, we are the first to demonstrate empirically how the social self can help distinguish between stereotype threat and priming effects. For instance, Wheeler and Petty (2001) have also detailed how the social self plays a role in stereotype threat, but they, unlike us, allowed the social self to play a smaller role than stereotype activation plays. We think the roles should be reversed, such that stereotype activation plays a supporting role to the leading role of the social self. Therein lies the fundamental difference between priming and stereotype threat. Stereotype threat is a situational predicament in which negative stereotypes are activated. Then, depending on who one is, concerns about confirming the negative stereotype may be increased if the stereotype is self-relevant. Although stereotype priming also relies on stereotype activation, the similarities stop there. Once the negative stereotype becomes linked to the social self and how a person performs, it is no longer a priming effect: It becomes a stereotype threat effect.

In short, we argue that stereotype threat involves knowing as well as being the stereotype. Thus, stereotype threat occurs for people who know and are targeted by the relevant stereotype, leading them to feel threatened ("I worry about my performance because I am a woman, and I know women are perceived to be bad at math"), whereas priming effects can occur for anyone because no such link is necessary. Therefore, when distinguishing between these two routes to stereotype-based performance (stereotype priming, stereotype threat), it seems critical to include targets and nontargets, stereotype threat conditions (i.e., when targets' performance is relevant to the group stereotype) and nonstereotype threat conditions (i.e., when targets' performance is irrelevant to the group stereotype), as well as measures of participants' threat-based concerns in the research design. By including all of these factors, we can assess how priming of negative traits and stereotype threat differentially affect stereotyped and nonstereotyped participants' test performance and threat-based concerns, something that prior research on negative trait priming was unable to do (e.g., Ambady et al., 2004; Oswald & Harvey, 2000; Wheeler et al., 2001).

Research Overview

In three experiments, we compared and contrasted priming with stereotype threat so that we could test the knowing-and-being logic of stereotype threat. The first two experiments pitted the effects of priming against stereotype threat on participants' math test (Experiment 1) and emotion test (Experiment 2) performance and their threat-based concerns. The final experiment was conducted to examine further our reasoning regarding the knowing-and-being logic of stereotype threat and how merely heightening the accessibility of a generalized social self can cause targets, but not nontargets, to have lower test performance and higher threat-based concerns as a result. In other words, we tested the notion that accessibility of the social self just before taking a nonthreatening test is sufficient for causing stereotype threat among targets, but not nontargets, because of the negative stereotype associated with targets' social self.

Experiment 1: Math Test Performance

The primary goal of Experiment 1 was to highlight the distinction between priming and stereotype threat. To do this, we focused on the stereotype that women possess less math ability than men. Accordingly, male and female participants took a difficult math test under stereotype threat or nonstereotype threat conditions. In addition, half the participants were primed with the negative trait *dumb* and its semantic associates before they took the test, whereas the other half were not primed.

On the basis of our knowing-and-being reasoning, we expected that we would find the typical stereotype threat pattern (i.e., targets having lower scores when the group stereotype is relevant to their performance) on targets' test performance and threat-based concerns in the stereotype threat conditions. We further predicted that in the nondiagnostic conditions both male and female participants would have lower scores in the primed than in the not primed condition but that there would be no difference in their threat-based concerns between these conditions, because priming should not affect threat-based concerns. In other words, we should find a general priming effect only on participants' test performance, but not on their threat-based concerns.

Method

Participants and Design

Participants were 60 female and 50 male Dutch undergraduates who took part in exchange for course credit. For this experiment, we used a 2 (gender of participant: female, male) × 2 (test description: diagnostic, nondiagnostic) × 2 (stereotype prime: not primed, primed) between-participants design.

Procedure

On entering the laboratory, participants were informed that they would be involved in a series of short, unrelated tasks.

Scrambled sentence task. Participants in the primed condition unscrambled sentences that were designed to activate the trait *dumb* or its semantic associate (e.g., *unintelligent*, *foolish*). To do this, participants were given a list of 20 scrambled five-word groups. Participants were then instructed to reorganize and write out the word groups into meaningful sentences, using at least four words from each group (cf. Srull & Wyer, 1979). Twelve of the word groups contained a concept that was relevant to the focal trait (*dumb*). Filler word groups containing behaviors not related to the trait (e.g., "a packed trip suitcase for") were interspersed among the word groups containing the priming stimuli. The participants in the not primed conditions completed the same number of sentences with only the filler word groups.

Test description manipulation. For all conditions, the test format resembled a standard Graduate Record Exam (GRE) math section but varied in a number of important ways. In the diagnostic condition, the test was described as being diagnostic of math ability as well as one that can identify a person's mathematical strengths and weaknesses. Moreover, written on the cover of the test booklet was the name of a fictitious testing

Table 1
Means and Standard Deviations of Math Test Performance and Threat-Based Concerns as a Function of Stereotype Prime, Test Description, and Participant Gender

Test description and gender	Stereotype prime							
	Not primed				Primed			
	Diagnostic		Nondiagnostic		Diagnostic		Nondiagnostic	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Math test performance								
Female participant	10.67_{a}	1.97	12.29_{b}	1.57	9.27_{d}	1.79	10.63_{a}	1.54
Male participant	13.50_{c}	1.45	12.25_{b}	0.87	12.50_{b}	1.45	10.86_{a}	1.17
Threat-based concerns								
Female participant	3.03_{b}	0.58	2.08_{a}	0.53	2.87_{b}	0.63	2.02_{a}	0.66
Male participant	2.31_{a}	0.56	1.92_{a}	0.38	1.78_{a}	0.64	2.21_{a}	0.61

Note. All means that do not share a common subscript differ at $p < .05$.

center, "Massachusetts Aptitude Assessment Center (MAAC)," followed by the label "Diagnostic Exam." This procedure has successfully created a situation of stereotype threat in previous research (Gonzales et al., 2002; Marx et al., 2005; Steele & Aronson, 1995). In the nondiagnostic condition, the same test was described as a reasoning exercise, thus purposefully not activating the negative stereotype about women and math. Furthermore, in both conditions, participants were told that they would receive feedback about their test performance at the conclusion of the experiment (no feedback was actually given). Participants had 30 min to complete the 20-problem math test. Performance could range from 0 to 20.

Threat-based concerns. To assess whether our manipulations affected the participants' threat-based concerns, we had them indicate how much they agreed with the following three statements: "I worry that my ability to perform well on math tests is affected by my gender"; "I worry that if I perform poorly on this test, the experimenter will attribute my poor performance to my gender"; "I worry that, because I know the negative stereotype about women and math, my anxiety about confirming that stereotype will negatively influence how I perform on math tests." Responses were recorded on a 7-point scale anchored with the terms (1) *strongly disagree* and (7) *strongly agree*. We averaged the participants' responses to form a single threat-based concern score ($\alpha = .74$).[2] When participants were finished, they were debriefed and thanked for their time.

Results and Discussion

Math Test Performance

The participants' math test performance was analyzed by means of a 2 (participant gender) × 2 (test description) × 2 (stereotype prime) analysis of variance (ANOVA; see Table 1). This analysis revealed main effects for participant gender, $F(1, 102) = 28.67$, $p < .05$, $\eta^2 = .22$, and stereotype prime, $F(1, 102) = 21.86$, $p < .05$, $\eta^2 = .18$. We also found a Participant Gender × Test Description interaction, $F(1, 102) = 25.32$, $p < .05$, $\eta^2 = .20$ (all other effects, Fs < 1.00).

Table 1 shows, as hypothesized, that within the primed–diagnostic condition, female participants ($M = 9.27$, $SD = 1.79$) underperformed relative to male participants ($M = 12.50$, $SD = 1.45$), $F(1, 102) = 30.21$, $p < .05$, $\eta^2 = .23$. This same pattern occurred for female ($M = 10.67$, $SD = 1.97$) and male ($M = 13.50$, $SD = 1.45$) participants in the not primed–diagnostic condition, $F(1, 102) = 20.87$, $p < .05$, $\eta^2 = .17$. We also found that female participants in the not primed–diagnostic condition ($M = 10.67$, $SD = 1.97$) had lower math scores than female participants in the not primed–nondiagnostic condition ($M = 12.29$, $SD = 1.57$), $F(1, 102) = 8.02$, $p < .05$, $\eta^2 = .07$.[3] On the whole, these results demonstrate the standard stereotype threat effect (i.e., targets performing poorly when a group stereotype is relevant to their performance).

Our results showed that male participants in the primed–nondiagnostic condition ($M = 10.86$, $SD = 1.17$) had lower math scores than male participants in the not primed–nondiagnostic condition ($M = 12.25$, $SD = 0.87$), $F(1, 102) = 5.43$, $p < .05$, $\eta^2 = .05$. This pattern also occurred in the nondiagnostic condition for female participants in the primed ($M = 10.63$, $SD = 1.54$) compared with the not primed conditions ($M = 12.29$, $SD = 1.57$), $F(1, 102) = 9.88$, $p < .05$, $\eta^2 = .09$. Taken together, these results are consistent with a general priming effect. These results also indicate that when the social self is not accessible, priming harms targets' test performance in the same way that it does nontargets.

Interestingly, we also found a lift effect (Walton & Cohen, 2003) in the primed condition, such that male participants performed better in the diagnostic condition ($M = 12.50$, $SD = 1.45$) than in the nondiagnostic condition ($M = 10.86$, $SD = 1.17$), $F(1, 102) = 7.55$, $p < .05$, $\eta^2 = .07$. Male participants in the not primed condition also performed better in the diagnostic ($M = 13.50$, $SD = 1.45$) relative to the nondiagnostic condition ($M = 12.25$, $SD = 0.87$), $F(1, 102) = 4.07$, $p < .05$, $\eta^2 = .04$. Next, we turned to the question of whether priming has an effect on targets'

[2] It is important to note that, as in all the experiments reported in this article, separate analyses for each of the threat-based concern items yielded a similar pattern of effects. For reasons of parsimony, we only reported the effects on the composite scores.

[3] For Experiments 1 and 2, we did not compare targets' test performance between the diagnostic and nondiagnostic conditions within the primed condition because targets' test performance should be negatively affected in the primed–nondiagnostic condition (as would be expected from a priming perspective); thus, their test performance may not differ reliably from targets' test performance in the diagnostic condition. Moreover, given that priming should negatively affect targets' performance in the nondiagnostic condition, it is unclear what the comparison between the diagnostic and nondiagnostic condition would mean within the primed condition.

threat-based concerns, as this is also a critical test of the knowing-and-being logic of stereotype threat (Marx & Stapel, in press; Marx et al., 2005; Steele, 1997; Steele et al., 2002).

Threat-Based Concerns

Participants' threat-based concerns were analyzed by means of a 2 (participant gender) × 2 (test description) × 2 (stereotype prime) ANOVA (see Table 1). This analysis revealed main effects for participant gender, $F(1, 102) = 13.09$, $p < .05$, $\eta^2 = .12$, and test description, $F(1, 102) = 17.62$, $p < .05$, $\eta^2 = .14$. There was also a Participant Gender × Test Description interaction, $F(1, 102) = 17.47$, $p < .05$, $\eta^2 = .14$, and a marginally reliable Test Description × Stereotype Prime interaction, $F(1, 102) = 3.77$, $p < .06$, $\eta^2 = .04$ (all other effects, ps $> .11$).

As can be seen in Table 1, within the primed–diagnostic condition female participants ($M = 2.87$, $SD = 0.63$) experienced more concern relative to male participants ($M = 1.78$, $SD = 0.64$), $F(1, 102) = 23.36$, $p < .05$, $\eta^2 = .18$. Within the not primed–diagnostic condition, we also found that female participants ($M = 3.03$, $SD = 0.58$) had higher concern scores than male participants ($M = 2.31$, $SD = 0.56$), $F(1, 102) = 9.18$, $p < .05$, $\eta^2 = .08$. Moreover, female participants in the primed condition had higher concern scores in the diagnostic ($M = 2.87$, $SD = 0.63$) than in the nondiagnostic conditions ($M = 2.02$, $SD = 0.66$), $F(1, 102) = 16.50$, $p < .05$, $\eta^2 = .14$. Female participants in the not primed condition likewise experienced more concern in the diagnostic ($M = 3.03$, $SD = 0.58$) than in the nondiagnostic condition ($M = 2.08$, $SD = 0.53$), $F(1, 102) = 18.73$, $p < .05$, $\eta^2 = .15$. In the diagnostic condition, there was no reliable difference in female participants' threat-based concern between the primed ($M = 2.87$, $SD = 0.63$) and not primed conditions ($M = 3.03$, $SD = 0.58$, $F < 1.00$). Furthermore, there were no differences between the female and male participants' threat-based concern scores in the nondiagnostic conditions (ps $> .20$). Together, these results demonstrate the typical stereotype threat pattern on targets' threat-based concerns and show that priming does not have any meaningful effect on their threat appraisals.

In short, our findings highlight how stereotype threat and negative trait priming effects differ. In the diagnostic condition, female participants underperformed compared with male participants. Results also showed that female participants had lower math scores in the diagnostic condition relative to female participants in the nondiagnostic condition. Moreover, in the nondiagnostic condition we found the typical priming effect on participants' math test performance (stereotype activation → stereotype-consistent test performance), such that both male and female participants performed in a stereotype-consistent manner. It is important that, as predicted, the prime did not elevate female participants' threat-based concerns. These results support the notion that the activation of stereotypic traits is not enough for stereotype threat to occur. For stereotype threat to occur, targets need to make the connection between the stereotype and how well they perform in a specific testing situation; hence, in line with the knowing-and-being logic of stereotype threat, they need to link what they know about the stereotype to who they are.

Experiment 2: Emotion Test Performance

The purpose of Experiment 2 was to replicate the effects of Experiment 1 as well as to generalize our knowing-and-being logic of stereotype threat effects to domains other than academics (e.g., Aronson et al., 1999; Leyens, Desert, Croizet, & Darcis, 2000; Marx & Stapel, in press). We also examined whether threat and trait priming lead to the same level of endorsement regarding the negative stereotype about men and emotional insensitivity, as it could be reasoned that participants' stereotype endorsement may contribute to possible performance differences (cf. Blanton, Christie, & Dye, 2002). Inclusion of this item also serves as a manipulation check of sorts, in that we could assess whether there was equal endorsement of the stereotype as well as whether the participants were even aware of the stereotype about men and emotional insensitivity. This issue seemed important given that we were focusing on a relatively less studied stereotype than the stereotypes used in past work on stereotype threat and priming (e.g., women and math).

To investigate these issues, we used procedures similar to the ones used in Experiment 1 except that participants took an emotion test and were primed with the negative trait *emotional insensitivity* and its semantic associates. We made the same predictions as in Experiment 1, but in this experiment male participants were the targets of a negative stereotype.

Method

Participants and Design

Participants were 54 female and 51 male Dutch undergraduates who took part in exchange for course credit. For this experiment we used a 2 (gender of participant: female, male) × 2 (test description: diagnostic, nondiagnostic) × 2 (stereotype prime: not primed, primed) between-participants design.

Procedure

On entering the laboratory, participants were informed that they would be involved in a series of short, unrelated tasks.

Priming manipulation. We used the scrambled sentence task from Experiment 1 but modified it so that the priming stimuli activated the trait *emotional insensitivity* as well as its semantic associates (e.g., *inconsiderate, cold*). The filler word groups were identical to those used in Experiment 1.

Emotion test. We used the same test description manipulation and test format from Experiment 1, but for this experiment we adapted them so that they fit with the stereotype about men and emotional insensitivity. Moreover, written on the cover of the diagnostic test booklet was the label "Emotional Sensitivity Exam" and written on the cover of the nondiagnostic test booklet was the label "Emotional Exercise." The emotion test comprised several types of problems that were loosely based on other emotion measures and exercises (see Bar-On, 1997; Schutte et al., 1998). For instance, participants had to answer problems about which emotion is best captured by a particular facial expression, indicate which two basic emotions (e.g., joy, expectation) make a more complex emotion (optimism), and answer problems about how emotions typically develop (e.g., "If you feel more and more guilty and you lose your feeling of self-worth then you feel?"; shame). A version of this emotion test has been used in our past work on emotional stereotype threat (see for details, Marx & Stapel, in press). Participants had 20 minutes to complete the 10-problem emotion test. Test performance could range from 0 to 10.

Table 2
Means and Standard Deviations of Emotion Test Performance and Threat-Based Concerns as a Function of Stereotype Prime, Test Description, and Participant Gender

Test description and gender	Stereotype prime							
	Not primed				Primed			
	Diagnostic		Nondiagnostic		Diagnostic		Nondiagnostic	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Emotion test performance								
Female participant	8.92_{a}	1.00	6.64_{b}	1.21	6.92_{b}	1.38	5.28_{c}	1.07
Male participant	5.08_{c}	1.38	7.17_{b}	0.94	$4.07_{d,c}$	1.86	4.83_{c}	1.03
Threat-based concerns								
Female participant	1.56_{a}	0.62	1.64_{a}	1.04	1.90_{a}	1.37	1.78_{a}	1.12
Male participant	2.67_{b}	0.84	1.39_{a}	0.49	3.31_{b}	0.81	1.72_{a}	1.05

Note. All means that do not share a common subscript differ at $p < .05$.

Threat-based concerns and stereotype endorsement. To assess participants' threat-based concerns, we modified the concern measure from Experiment 1 so that it was appropriate for the stereotype about men and emotional insensitivity ($\alpha = .72$). After completing the threat-based concern items, participants answered an item about their endorsement of the stereotype about men and emotional insensitivity ("Women are more emotionally sensitive than men.").[4] Scores on this item could range from 1 (*not at all*) to 9 (*very much*), with higher numbers indicating more endorsement of the stereotype. Threat-based concerns and the stereotype endorsement item were not strongly related ($r = .12$, $p = .22$). On completion of these items, the participants were debriefed and thanked for their time.

Results and Discussion

Stereotype Endorsement

We analyzed participants' stereotype endorsement scores using a 2 (participant gender) × 2 (test description) × 2 (stereotype prime) ANOVA. We found, as anticipated, no main or interactive effects (*p*s > .27). That we found no differences suggests that neither our threat nor our priming manipulations affected participants' level of endorsement regarding the stereotype about men and emotional insensitivity. It is also clear that participants in this experiment endorsed the stereotype equally. Hence, if we find effects as a function of our manipulations, it would be unlikely that those effects were due to differences in participants' endorsement of the stereotype.[5]

Emotion Test Performance

The participants' emotion test performance was analyzed by means of a 2 (participant gender) × 2 (test description) × 2 (stereotype prime) ANOVA (see Table 2). This analysis revealed main effects for participant gender, $F(1, 97) = 44.02$, $p < .05$, $\eta^2 = .31$, and stereotype prime, $F(1, 97) = 48.77$, $p < .05$, $\eta^2 = .34$. We also found a Participant Gender × Test Description interaction, $F(1, 97) = 43.86$, $p < .05$, $\eta^2 = .31$, and a marginally reliable three-way interaction, $F(1, 97) = 3.82$, $p < .06$, $\eta^2 = .04$ (all other effects, *p*s > .24).

Table 2 shows, as predicted, that in the primed–diagnostic condition female participants ($M = 6.92$, $SD = 1.38$) outperformed male participants ($M = 4.07$, $SD = 1.86$), $F(1, 97) = 33.80$, $p < .05$, $\eta^2 = .26$. In the not primed–diagnostic condition, we also found that female participants ($M = 8.92$, $SD = 1.00$) had higher scores than male participants ($M = 5.08$, $SD = 1.38$), $F(1, 97) = 56.80$, $p < .05$, $\eta^2 = .37$. Moreover, when we compared male participants' test performance within the not primed condition we found that they performed better in the nondiagnostic condition ($M = 7.17$, $SD = 0.94$) than in the diagnostic condition ($M = 5.08$, $SD = 1.38$), $F(1, 97) = 16.83$, $p < .05$, $\eta^2 = .15$. These comparisons again demonstrate the typical stereotype threat pattern, but this time on male participants' emotion test performance.

Within the nondiagnostic conditions, male participants had lower scores in the primed condition ($M = 4.83$, $SD = 1.03$) than in the not primed condition ($M = 7.17$, $SD = 0.94$), $F(1, 97) = 20.28$, $p < .05$, $\eta^2 = .17$. This was also the case for female participants in the primed ($M = 5.28$, $SD = 1.07$) and not primed conditions ($M = 6.64$, $SD = 1.21$), $F(1, 97) = 7.80$, $p < .05$, $\eta^2 = .07$. Again, these effects underscore our contention that when targets do not have to contend with a negative group stereotype, priming harms their test performance in the same way that it does nontargets.

Just as in Experiment 1 for math test performance, we found a lift effect for female participants in emotion test performance such

[4] Of course, we are aware that this statement specifically asks participants about their endorsement of the positive stereotype about women and emotional sensitivity. However, this statement also implies that men possess less emotional sensitivity than do women; thus, it serves as a reasonable proxy for endorsement of the stereotype about men and emotional insensitivity.

[5] In addition to examining whether our stereotype endorsement item was affected by our experimental manipulations, we also examined whether participants knew the stereotype about men and emotional insensitivity. To do this, we looked at the percentage of participants who had scores above the midpoint of our scale (5), as this would indicate that participants were aware of the stereotype, whereas scores below the midpoint would indicate that participants were unaware of the stereotype. We found that 4% of participants were below the midpoint, 12% were at the midpoint, and 84% were above the midpoint.

that in the primed–diagnostic conditions ($M = 6.92$, $SD = 1.38$) female participants performed better than female participants in the primed–nondiagnostic conditions ($M = 5.28$, $SD = 1.07$), $F(1, 97) = 12.53$, $p < .05$, $\eta^2 = .12$. The same effect occurred in the not primed condition for female participants in the diagnostic ($M = 8.92$, $SD = 1.00$) and nondiagnostic conditions ($M = 6.64$, $SD = 1.21$), $F(1, 97) = 18.42$, $p < .05$, $\eta^2 = .16$. Next, we turned to the question of whether targets' threat-based concerns are affected by priming of stereotypic traits.

Threat-Based Concerns

The participants' threat-based concerns were analyzed by means of a 2 (participant gender) × 2 (test description) × 2 (stereotype prime) ANOVA (see Table 2). We found main effects for participant gender, $F(1, 97) = 9.03$, $p < .05$, $\eta^2 = .08$, test description, $F(1, 97) = 14.64$, $p < .05$, $\eta^2 = .13$, and stereotype prime, $F(1, 97) = 4.59$, $p < .05$, $\eta^2 = .04$. We also found a Participant Gender × Test Description interaction, $F(1, 97) = 13.90$, $p < .05$, $\eta^2 = .12$ (all other effects, Fs < 1.00).

Table 2 shows that within the primed–diagnostic condition, male participants ($M = 3.31$, $SD = 0.81$) had higher concern scores relative to female participants ($M = 1.90$, $SD = 1.37$), $F(1, 97) = 14.33$, $p < .05$, $\eta^2 = .13$. The same effect occurred in the not primed–diagnostic condition such that male participants ($M = 2.67$, $SD = 0.84$) were more concerned than female participants ($M = 1.56$, $SD = 0.62$), $F(1, 97) = 8.22$, $p < .05$, $\eta^2 = .08$. We also found that male participants in the primed–diagnostic condition ($M = 3.31$, $SD = 0.81$) felt more concern than male participants in the primed–nondiagnostic conditions ($M = 1.72$, $SD = 1.05$), $F(1, 97) = 17.47$, $p < .05$, $\eta^2 = .15$. This was also the case for male participants in the not primed–diagnostic conditions ($M = 2.67$, $SD = 0.84$) compared with male participants in the not primed–nondiagnostic condition ($M = 1.39$, $SD = 0.49$), $F(1, 97) = 10.93$, $p < .05$, $\eta^2 = .10$. Within the diagnostic conditions, there was a marginal difference in male participants' threat-based concern scores between the primed ($M = 3.31$, $SD = 0.81$) and not primed conditions ($M = 2.67$, $SD = 0.84$), $F(1, 97) = 2.95$, $p = .09$, $\eta^2 = .03$, indicating that our priming procedure had a slight effect on the male participants' concern scores. It is important that, as would be predicted from stereotype threat theory, there were no differences between male and female participants' threat-based concern scores in the nondiagnostic conditions, Fs < 1.00.

Taken together, the findings from Experiment 2 provide additional support for how stereotype threat and priming effects differ. As before, we found the standard stereotype threat pattern on participants' test performance. In the nondiagnostic condition, we found the typical priming effect (stereotype activation → stereotype-consistent test performance), such that both targets and nontargets performed in a stereotype-confirming manner. But what is probably more critical to our reasoning regarding the knowing-and-being aspect of stereotype threat is the fact that the prime in the nondiagnostic condition did not elevate targets' threat appraisals nor did our manipulations lead to differences in participants' stereotype endorsement.

In sum then, these results show that activation of stereotypic traits is not sufficient for targets to make the connection between the stereotype and their performance in the testing situation: Participants need to know as well as be the stereotype in order for the stereotype to affect their threat-based concerns (Major & O'Brien, 2005; Marx et al., 1999, 2005; Steele et al., 2002). This reasoning is also supported by the fact that when we correlated participants' emotion test performance with their stereotype endorsement and threat-based concern scores, controlling for the experimental variables (i.e., gender, test description, priming), we only found a reliable partial correlation between test performance and threat-based concerns ($r = -.20$, $p < .04$). The partial correlation between test performance and stereotype endorsement was not reliable ($r = -.02$, $p = .83$); thus, stereotype endorsement and threat-based concerns may co-occur, but only threat-based concerns should be affected by being the stereotype. Simply knowing the stereotype is not enough to cause stereotype threat; targets need to link being the stereotype to how they perform in the situation.

Experiment 3: The Social Self and Stereotype Threat

Now that we have demonstrated that priming and stereotype threat yield a different pattern of effects on participants' test performance and threat-based concerns, we turned to the question of whether increasing accessibility of the social self just prior to taking a test would create a stereotype threat experience for targets, but not nontargets, even when the test is nonthreatening in nature (when it is presented as a reasoning exercise). This notion is based on our prior work showing that stereotype threat leads to greater accessibility of targets' stereotyped identity compared with other equally important social identities (e.g., friends and family; Marx et al., 2005). Thus, we argue that it should be possible to cause poor performance among targets, not only by using a standard stereotype threat manipulation (i.e., test diagnosticity) but also by making the social self accessible before taking a nonthreatening test. Of course, activating a general social self (we-ness) may make many social identities accessible, but once the context is factored in (e.g., I am a person taking a math test) then the social identity (gender) that is most relevant to the situation (taking a math test) may win out in the end (e.g., Major & O'Brien, 2005; Onorato & Turner, 2004; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). And if this social identity is linked to a negative stereotype, then it seems reasonable to suggest that one's threat-based concerns may also be affected (e.g., Marx et al., 1999, 2005).

Although our manipulation of the social self is akin to manipulations used in previous research on stereotype-based performance (see Shih et al., 1999; Steele & Aronson, 1995), our approach deviates from this past work in that we do not ask participants to indicate their stereotyped identity prior to taking a test. Thus, we do not purposefully remind targets about the associated group stereotype. This approach therefore allows us to highlight the critical role of the situation and how the situational cues may actually shape a general awareness of one's social self into a specific awareness of a particular social identity—the identity that is most closely aligned with the situation and associated stereotype (e.g., Onorato & Turner, 2004; Turner et al., 1987; see also Major & O'Brien, 2005; Marx et al., 1999, for discussions of contextual effects in stereotype threat).

To test this line of reasoning regarding the interaction of the social self and the situation, we asked male and female participants to take a math test under one of four conditions. In the first condition, we increased the accessibility of participants' social self before they took a nonthreatening math test. In the second condi-

tion, we primed participants with stereotypic traits (see Experiment 1) before they took a nonthreatening math test. In the third condition, the test was described as diagnostic of math ability (stereotype threat), and in the fourth condition participants simply took a nonthreatening test (no stereotype threat). By including these four conditions, we were able to accomplish two things. One, we could replicate the effects from Experiments 1 and 2, and two, we could highlight how knowing the stereotype (e.g., "women are bad at math") and being from the stereotyped group (e.g., "I am a woman") in combination with the situational cues can lead to higher threat-based concerns and lower test performance among targets, but not nontargets.

For this last experiment, we expected to find the typical stereotype threat pattern on targets' math test performance and threat-based concerns in the diagnostic and nondiagnostic conditions. We also anticipated that female participants would have lower scores and higher threat-based concerns than male participants in the social self–nondiagnostic condition but that they would perform equally as poorly and not differ in their threat-based concerns in the *dumb* prime–nondiagnostic condition. We further predicted that male and female participants' test performance would be lower in the *dumb* prime–nondiagnostic conditions than in the nondiagnostic conditions but that there would be no difference in their threat-based concern scores. And finally, we hypothesized that the pattern of performance and threat-based concerns would be the same in both the social self–nondiagnostic and the diagnostic conditions. This predicted pattern serves as the critical test of our reasoning regarding the knowing-and-being aspect of stereotype threat, as it would show that underperformance can occur for targets via the combination of social self activation and the situational cues or via more traditional stereotype threat manipulations (test diagnosticity). That is, both manipulations make those aspects of the social self that are most associated with the stereotype particularly accessible at the time, thus leading to underperformance for targets, but not nontargets.

Method

Participants and Design

Participants were 57 female and 46 male Dutch undergraduates who took part in exchange for course credit. For this experiment, we used a 2 (gender of participant: female, male) × 4 (type of condition: social self–nondiagnostic test, *dumb* prime–nondiagnostic test, diagnostic test, nondiagnostic test) between-participants design.

Procedure

On entering the laboratory, participants were informed that they would be involved in a series of short, unrelated tasks. One quarter of the participants completed a scrambled sentence task and then took a nondiagnostic math test. Another quarter of the participants completed a word search task and then took a nondiagnostic math test. Another quarter took a diagnostic math test, and the final quarter took a nondiagnostic math test.

Scrambled sentence task. We used the scrambled sentence task from Experiment 1.

Word search task. For this task, participants were told that as part of a proofreading and word search exercise they would read a short paragraph detailing a trip to the city. They were further told that as they read the paragraph they should circle all the pronouns that appeared in the text. The pronouns in the text were *we, our, ourselves,* and *us* (see for details Brewer & Gardner, 1996). Past research has shown that this procedure is effective in increasing feelings of we-ness or accessibility of the social self (Brewer & Gardner, 1996; Stapel & Koomen, 2001).

Test description manipulation. We used the test description manipulation (diagnostic math test or reasoning exercise) and math test from Experiment 1.

Threat-based concerns. We used the threat-based concerns measure from Experiment 1 ($\alpha = 74$). On completion of this measure, participants were debriefed and thanked.

Results and Discussion

Math Test Performance

The participants' math test performance was analyzed by means of a 2 (participant gender) × 4 (condition type) ANOVA (see Table 3). We found main effects for participant gender, $F(1, 95) = 9.20$, $p < .05$, $\eta^2 = .09$, and condition type, $F(3, 95) = 6.40$, $p < .05$, $\eta^2 = .06$. We also found an omnibus interaction, $F(3, 95) = 2.92$, $p < .05$, $\eta^2 = .08$.

Table 3 shows that within the diagnostic condition female participants ($M = 10.54$, $SD = 0.97$) underperformed relative to male participants ($M = 12.31$, $SD = 1.97$), $F(1, 95) = 8.88$, $p < .05$, $\eta^2 = .09$. Moreover, female participants in the diagnostic condition ($M = 10.54$, $SD = 0.97$) had lower scores than female participants in the nondiagnostic condition ($M = 12.47$, $SD = 1.64$), $F(1, 95) = 11.33$, $p < .05$, $\eta^2 = .11$. As expected, there was

Table 3
Means and Standard Deviations of Math Test Performance and Threat-Based Concerns as a Function of Condition Type and Participant Gender

Test description and gender	Condition type							
	Diagnostic		Nondiagnostic		Social self		Dumb prime	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Math test performance								
Female participants	10.54_{a}	0.97	12.47_{b}	1.64	11.07_{a}	1.90	10.67_{a}	1.76
Male participants	12.31_{b}	1.97	12.45_{b}	0.82	12.91_{b}	0.70	10.73_{a}	1.42
Threat-based concerns								
Female participants	2.87_{b}	0.52	1.93_{a}	0.47	2.76_{b}	0.76	1.93_{a}	0.63
Male participants	1.90_{a}	0.53	1.85_{a}	0.31	$1.61_{a,c}$	0.47	2.06_{a}	0.53

Note. All means that do not share a common subscript differ at $p < .05$.

no difference between male and female participants' test performance in the nondiagnostic conditions ($F < 1.00$). These results are thus consistent with a typical stereotype threat effect. Critical to our reasoning about accessibility of the social self and the contextual cues, we found that in the social self–nondiagnostic condition female participants ($M = 11.07$, $SD = 1.90$) had lower scores than male participants ($M = 12.91$, $SD = 0.70$), $F(1, 95) = 9.10$, $p < .05$, $\eta^2 = .09$. This result is similar to what we found in the diagnostic condition, demonstrating that accessibility of the social self leads to the same effect as the test diagnosticity manipulation because the context helps shape a general awareness of the social self into a more specific awareness of the social self (gender) that is relevant to the situation (cf. Shih et al., 1999; Steele & Aronson, 1995; see also Onorato & Turner, 2004; Turner et al., 1987). Our final comparison showed that male and female participants had lower math scores in the *dumb* prime–nondiagnostic condition than in the nondiagnostic condition, $F(1, 95) = 17.15$, $p < .05$, $\eta^2 = .15$, thus underscoring our point that when the group stereotype is not linked to how participants perform, priming has a similar effect on targets' and nontargets' performance. We next assessed whether our manipulations affected participants' threat-based concerns, as this measure could further show that targets are linking how they perform on the test to the stereotype associated with their social self.

Threat-Based Concerns

We analyzed the participants' threat-based concerns using a 2 (participant gender) × 4 (condition type) ANOVA (see Table 3). There were main effects for participant gender, $F(1, 95) = 25.54$, $p < .05$, $\eta^2 = .21$, and condition type, $F(3, 95) = 5.71$, $p < .05$, $\eta^2 = .06$. We also found an omnibus interaction, $F(3, 95) = 9.55$, $p < .05$, $\eta^2 = .23$.

Table 3 shows that in the diagnostic condition, female participants ($M = 2.87$, $SD = 0.52$) had higher concern scores compared with male participants ($M = 1.90$, $SD = 0.53$), $F(1, 95) = 22.82$, $p < .05$, $\eta^2 = .19$. This effect highlights our main contention that when the stereotype (knowing) is relevant to one's social self (being), targets' threat-based concerns are elevated compared with nontargets. The same effect occurred in the social self–nondiagnostic condition, such that female participants ($M = 2.76$, $SD = 0.76$) had higher concern scores relative to male participants ($M = 1.61$, $SD = 0.47$), $F(1, 95) = 30.40$, $p < .05$, $\eta^2 = .24$. This last result provides additional support for the notion that when a stereotype is relevant to targets' test performance, either via the test diagnosticity manipulation or by increasing accessibility of the social self before taking a nonthreatening test, targets experience more concern than when the stereotype is irrelevant to their performance or when their social self is not as accessible. There were no differences, as expected, in participants' threat-based concerns between the nondiagnostic and *dumb* prime–nondiagnostic conditions (Fs < 1.00).

On the whole, these analyses make it quite apparent that when the stereotype is relevant to the social self (either because the test is presented as diagnostic or because participants' social self is cognitively accessible) only stereotyped targets underperform, but when the stereotype is irrelevant to how participants perform (i.e., the nondiagnostic condition) both targets and nontargets react in a corresponding manner to the situation: When primed with stereotypic traits, they underperform; when not primed with such traits, they perform better. It is important that trait priming did not influence targets' threat-based concerns, thus highlighting the link between accessibility of the social self and targets' concerns about confirming a negative stereotype that is associated with their group.

General Discussion

The three experiments presented here show how to distinguish between priming and stereotype threat effects. Throughout this article, we have argued that even though the outcomes of stereotype threat and priming often look the same (stereotype confirming poor performance), they are in fact different. Stressing this distinction seemed particularly important in light of the literature on stereotype-based performance, which has made the assumption that mere negative stereotype activation can lead to stereotype threat (Ambady et al., 2004; Dijksterhuis & Bargh, 2001; Dijksterhuis & Corneille, 2004; Gladwell, 2005; Oswald & Harvey, 2000; Wheeler et al., 2001; Wheeler & Petty, 2001). It is interesting that despite this growing body of research, no direct empirical comparison of stereotype threat and negative trait priming had been made until now. Moreover, we believe that it is important to demonstrate the distinction between these accounts so that there is a better understanding of stereotype-based underperformance in general.

As the present research shows, participants' threat-based concerns make the difference between priming and stereotype threat effects quite clear. Indeed, we found very consistent results using two types of tests (a math test and an emotions test) and stereotypes (i.e., "women are bad at math," "men are emotionally insensitive"). Taken together, the results from the first two experiments demonstrate that stereotype threat adversely affects targets', but not nontargets', test performance and threat-based concerns because only targets' social self is linked to the negative group stereotype. For example, in stereotype threat situations, targets (but not nontargets) are affected because they know the group-based stereotype ("women are bad at math"; "men are emotionally insensitive") and because they are members of the group that is targeted by the stereotype ("I am a woman"; "I am a man"). This issue of knowing-and-being the stereotype, however, has received little to no attention in research examining the relationship between priming of negative traits and stereotype threat (e.g., Ambady et al., 2004; Dijksterhuis & Corneille, 2004; Wheeler et al., 2001; but see Shih et al., 2002, for work on priming positive traits). In the nondiagnostic condition, we found the standard priming effect (stereotype activation → stereotype-consistent poor test performance), such that both targets and nontargets performed in a stereotype-confirming manner. But what is probably most important to our reasoning regarding the knowing-and-being aspect of stereotype threat is the fact that the prime did not elevate targets' concern scores in either the diagnostic or nondiagnostic conditions. Our final experiment revealed strong support for the notion that when the social self is accessible, only stereotyped targets underperform because accessibility of the social self is also linked to targets', but not nontargets', threat-based concerns (see Marx et al., 2005). This is not the case when targets are primed with negative traits. Simple trait priming procedures do not activate

such a link; thus, they have comparable effects on targets' and nontargets' test performance.

In summary, the findings from these experiments not only increase our understanding of stereotype threat, they also advance stereotype threat theory by showing that accessibility of the social self in a challenging testing situation is associated with the same threat-based concerns that accompany more typical manipulations of stereotype threat (i.e., test diagnosticity). These results also have several implications for stereotype threat and priming research.

Implications for Priming Effects

It is widely accepted that priming effects are rather straightforward, such that perceiving leads to behaving. And for the most part, this may be an accurate portrayal of priming effects, but as recent research has shown, the conditions for this priming effect may not be that simple. For example, Spears, Gordijn, Dijksterhuis, and Stapel (2004) have demonstrated that when group differences are salient, the influence of the prime works differently from the usual perception-to-behavior effect for those participants who are not targeted by the stereotype. That is, when participants are aware of their group membership, they have a tendency to distance themselves from stereotypes that are not relevant for their group: College students in this case would walk faster rather than slower down a hallway when primed with traits associated with the stereotype of an older person, whereas older people would walk slower down the hallway when primed with the same traits. As Spears et al. (2004) argue, their research suggests that priming effects are not always unmediated, unidirectional effects: seeing often leads to doing, but sometimes seeing leads to doing the opposite, such as when group memberships are relevant to the seeing–doing sequence. In other words, when trying to understand priming effects in social contexts, it is important to take into account the notion that priming and behavioral outcomes seldom occur in a social vacuum. Perceivers often relate to the primed information in a certain way. Perceivers may view the primed information as in-group or out-group information, as stereotypical or nonstereotypical information, and this may dramatically affect the outcome of priming effects (see Spears et al., 2004).

This line of reasoning also fits well with the present research. For instance, we showed that in the nondiagnostic conditions when participants are not particularly aware of their social identity, they behaved in a stereotype-consistent manner after being primed with negative traits. However, when participants' social self (being) was accessible, then the situation (seeing math problems) triggered the negative group stereotype (knowing) for targets such that they performed in a stereotype-consistent manner, even though the test was described in nonthreatening terms. This did not occur for nontargets because they did not have to contend with a negative stereotype, nor is the stereotype particularly relevant to how they performed. As we argued earlier, this result came about because in stereotype threat situations, targets are worried about confirming a negative stereotype about their group if they underperform. Thus, we believe that the present findings demonstrate that general priming effects (Bargh et al., 1996; Dijksterhuis & van Knippenberg, 1998; Kawakami, Young, & Dovidio, 2002; Levy, 1996) may not be as straightforward as they seem. That is, simple priming explanations may only be appropriate descriptions of what is happening in those situations in which participants are unable to draw the connection between their behavior and the stereotype. In light of this, we argue that priming experiments cannot and should not be used as evidence for stereotype threat, as these experiments bypass one of the hallmarks of stereotype threat: the worry about confirming a negative stereotype associated with one's group (Major & O'Brien, 2005; Marx et al., 2005; Steele et al., 2002).

Implications for Stereotype Threat

As we have argued throughout this article, knowledge of a group stereotype is necessary but not sufficient for causing stereotype threat; both knowing and being are critical (see Marx & Stapel, in press; Marx et al., 2005; Steele, 1997; Steele et al., 2002). This knowing-and-being logic suggests that threatening testing situations are especially likely to increase the accessibility of the social identity that is most relevant to the stereotype. In light of this, we demonstrated that increased accessibility of the social self in combination with relevant situational cues (i.e., seeing math problems) could change a normal testing situation into a threatening one, such that targets underperform relative to nontargets (e.g., Major & O'Brien, 2005; Marx et al., 1999).

Moreover, we focused on what Steele and his colleagues (Marx et al., 1999; Steele, 1997; Steele et al., 2002) have argued is one of the core principles of the theory—the worry about being judged in terms of the stereotype associated with one's group—so that we could highlight the difference between stereotype threat and priming effects. Indeed, stereotype threat and trait priming may have the same effect on targets' performance, but they work via different routes.[6] When performance is not relevant to the group stereotype (a nonstereotype threat situation), targets and nontargets perform similarly after stereotype priming. But, when performance is relevant to the group stereotype (a stereotype threat situation), the prime does not change the typical stereotype threat pattern of effects. Given this, it seems reasonable to suggest that participants' threat-based concerns could differentiate between these two outcomes, as this measure shows whether participants are making the link between their social self and their performance in the testing situation. The present research also highlights the situational nature of stereotype threat by showing that cues in the testing session (e.g., seeing math problems in combination with increased accessibility of the social self) can undermine targets' test performance even when the test is described in neutral terms. This finding underscores the importance of the social self in stereotype threat (Major & O'Brien, 2005; Marx et al., 2005; see also Wheeler & Petty, 2001).

[6] It may be tempting here to argue for statistical mediation of our stereotype threat and priming effects. We believe that testing for mediation would be problematic in these experiments because the order in which we measured our variables prevents any strong causal arguments. Indeed, measuring participants' threat-based concerns prior to taking a test may substantially alter the testing situation (e.g., artificially raise concerns, or even diffuse the threat by allowing participants to vent their feelings) such that it would be difficult to make any tenable claims. Because of this, we attempted to highlight the mediating process of threat-based concerns on stereotype threat effects via experimentation. That is, we manipulated the testing situation so we could show that in stereotype threat situations targets' threat-based concerns are elevated, and in nonthreat situations their concerns are lower. We also showed that negative trait priming had no such effect on targets' threat-based concerns.

Separating the Present Work From Past Work on Stereotype-Based Performance

The last few years have seen a growing interest in the effects of stereotype activation on the performance of stereotyped and nonstereotyped targets (e.g., Ambady et al., 2004; Shih et al., 2002; Wheeler & Petty, 2001; Wheeler et al., 2001). In fact, some may be tempted to view the present work as simply replicating or extending this past work. We see clear differences, however. For example, the most obvious difference between the past work and the current work is that we focus on negative stereotypes, whereas other past research has focused on positive stereotypes (Shih et al., 2002). Furthermore, prior research examining negative trait priming (e.g., Ambady et al., 2004; Wheeler et al., 2001) has not included both stereotyped and nonstereotyped targets, nor have they included stereotype threat and nonthreat manipulations. And finally, none of the prior experiments that have looked at stereotype threat effects from a priming perspective have included a measure of worry or concern. In short, there are clear differences between the current work and past work on priming and threat.

From a theoretical perspective, there are a number of parallels between the present work and relevant past work, yet there are also a number of distinctions (see Wheeler & Petty, 2001). For example, in a recent review of the priming and threat literature, Wheeler and Petty (2001) suggested that stereotype threat could be due to the simple activation of stereotype-relevant concepts. We agree with this interpretation to some extent; however, our results clearly show that stereotype threat also involves the interaction of situational cues and a person's social identity. Therefore, stereotype activation is necessary, but it is not sufficient to cause stereotype threat. That is, once the stereotype becomes linked to the social self and how one performs in the situation, we argue that it is no longer a simple priming effect. Stereotype activation alone cannot cause stereotype threat, unless the activation likewise leads to an increase in a person's concerns about confirming that stereotype. In short, whereas previous discussions about the effects of stereotype activation and performance have been somewhat elusive about the relationship between priming and threat (e.g., Wheeler & Petty, 2001), we argue and demonstrate quite clearly that stereotype threat and priming effects both rely on stereotype activation, but once the social self and threat-based concerns are factored into the equation, the paths to poor performance diverge depending on who you are and how you view the stereotype (e.g., Major & O'Brien, 2005; Marx et al., 2005).

Coda

To date, the distinction between stereotype threat and priming effects has not been made explicitly nor has it been tested systematically; thus, some researchers have treated priming and stereotype threat as being one and the same (Ambady et al., 2004; Dijksterhuis & Bargh, 2001; Dijksterhuis & Corneille, 2004; Gladwell, 2005; Oswald & Harvey, 2000; Wheeler et al., 2001; Wheeler & Petty, 2001). Because of this, we felt it critical to show when and how these two accounts explain stereotype-based performance, particularly because many stereotype priming experiments often do not include all of the necessary factors (i.e., stereotype threat and nonstereotype manipulations, stereotyped and nonstereotyped targets, measures of threat-based concerns) to make empirical comparisons between stereotype threat and priming effects. In the current research, we included such factors and showed that priming of stereotypic traits can lead to underperformance for both targets and nontargets (when their social self is not involved), but it had no meaningful effect on targets' threat-based concerns. In sum, by highlighting how the social self and targets' threat-based concerns can distinguish between these two accounts of stereotype-based performance, our results add to the growing literature on stereotype threat as well as the already vast stereotype priming literature.

References

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology, 40,* 401–408.

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

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

Bar-On, R. (1997). *The emotional intelligence inventory (EQ-I). Technical manual.* Toronto, Ontario, Canada: Multi-Health Systems.

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

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self-representations. *Journal of Personality and Social Psychology, 71,* 83–93.

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.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology, 33,* 1–40.

Dijksterhuis, A., & Corneille, O. (2004). *On the relations between stereotype activation and intellectual underperformance.* Unpublished manuscript.

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

Gladwell, M. (2005). *Blink: The power of thinking without thinking.* New York: Little, Brown.

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.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11,* 365–371.

Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activation. *Personality and Social Psychology Bulletin, 28,* 3–15.

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

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

Major, B., & O'Brien, L. T. (2005). The social psychology of stigma. *Annual Review of Psychology, 56,* 393–421.

Marx, D. M., Brown, J. L., & Steele, C. M. (1999). Allport's legacy and the situational press of stereotypes. *Journal of Social Issues, 55,* 491–502.

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting female students' math test performance. *Personality and Social Psychology Bulletin, 28,* 1185–1197.

Marx, D. M., & Stapel, D. A. (in press). It depends on your perspective: The role of self-relevance in stereotype-based underperformance. *Journal of Experimental Social Psychology.*

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We can do it: The interplay of construal orientation and social comparisons under threat. *Journal of Personality and Social Psychology, 88,* 432–446.

Onorato, R. S., & Turner. J. C. (2004). Fluidity in the self-concept: The shift from personal to social identity. *European Journal of Social Psychology, 34,* 257–278.

Oswald, D. L., & Harvey, R. D. (2000). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social, 19,* 338–356.

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology, 40,* 152–168.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25,* 167–177.

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

Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology, 83,* 638–647.

Spears, R., Gordijn, E., Dijksterhuis, A., & Stapel, D. A. (2004). Reaction in action: Intergroup contrast in automatic behavior. *Personality and Social Psychology Bulletin, 30,* 605–616.

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

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660–1672.

Stapel, D. A., & Koomen, W. (2001). I, we, and the effects of others on me: How self-construal level moderates social comparison effects. *Journal of Personality and Social Psychology, 80,* 766–781.

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 vulnerability and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

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

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.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford, England: Blackwell.

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

Wheeler, S. C., Jarvis, B. G., & Petty, R. E. (2001). Think unto others: The self-destructive impact of negative racial stereotypes. *Journal of Experimental Social Psychology, 37,* 173–180.

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

Received May 9, 2005
Revision received November 24, 2005
Accepted November 28, 2005 ■