

Personality and Social Psychology Bulletin

<http://psp.sagepub.com>

Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally

Paul G. Davies, Steven J. Spencer, Diane M. Quinn and Rebecca Gerhardstein

Pers Soc Psychol Bull 2002; 28; 1615

DOI: 10.1177/014616702237644

The online version of this article can be found at:
<http://psp.sagepub.com/cgi/content/abstract/28/12/1615>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Society for Personality and Social Psychology](http://www.spsp.org)

Additional services and information for *Personality and Social Psychology Bulletin* can be found at:

Email Alerts: <http://psp.sagepub.com/cgi/alerts>

Subscriptions: <http://psp.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations <http://psp.sagepub.com/cgi/content/refs/28/12/1615>

Consuming Images: How Television Commercials That Elicit Stereotype Threat Can Restrain Women Academically and Professionally

Paul G. Davies
Stanford University

Steven J. Spencer
University of Waterloo

Diane M. Quinn
University of Connecticut

Rebecca Gerhardtstein
Florida State University

Women in quantitative fields risk being personally reduced to negative stereotypes that allege a sex-based math inability. This situational predicament, termed stereotype threat, can undermine women's performance and aspirations in all quantitative domains. Gender-stereotypic television commercials were employed in three studies to elicit the female stereotype among both men and women. Study 1 revealed that only women for whom the activated stereotype was self-relevant underperformed on a subsequent math test. Exposure to the stereotypic commercials led women taking an aptitude test in Study 2 to avoid math items in favor of verbal items. In Study 3, women who viewed the stereotypic commercials indicated less interest in educational/vocational options in which they were susceptible to stereotype threat (i.e., quantitative domains) and more interest in fields in which they were immune to stereotype threat (i.e., verbal domains).

Do you know of anything that is practiced by human beings in which the class of men doesn't excel that of women? Or shall we draw it out at length by speaking of weaving and the care of baked and boiled dishes—just those activities on which the reputation of the female sex is based and where its defeat is most ridiculous of all?

—Socrates (cited in Bloom, 1968)

They just don't know how to act with women students. They don't know what to do with you. Their whole attitude, and facial expressions and body language says, "You belong in the kitchen. What are you doing

here?" . . . There's a perceived threat that feeds male anger about bright women. People tell jokes that are negative to women, which just fuels it. And it encourages the women to pull back into safer ground—to choose more traditional majors, or to adopt a more neutral style—as a way of protecting themselves from that anger.

—Female science majors
(cited in Seymour & Hewitt, 1997)

Judging from the above quotes, stereotypes regarding the attributes of men and women have remained alarmingly stable over the past 25 centuries, which partially explains the clear gender divide in achievement-related choices that persists in our culture. For example, a striking gender disparity is found in the selection of college majors. Women, including those who are equally skilled and experienced in high school mathematics as their

Authors' Note: This article was based on a portion of Paul Davies's doctoral dissertation, which was conducted at the University of Waterloo. We are grateful to Ziva Kunda, Dov Cohen, Claude Steele, Duane Wegener, Mikki Hebl, Joshua Correll, and Michael Dahlin for their comments on an earlier version of this article. The research was supported by grants from the Social Sciences and Humanities Research Council of Canada and by doctoral scholarships from the Ontario Government. Correspondence concerning this article can be addressed to Paul G. Davies, Department of Psychology, Jordan Hall, Building 420, Stanford University, Stanford, CA 94305-2130; e-mail: pgdavies@psych.stanford.edu.

PSPB, Vol. 28 No. 12, December 2002 1615-1628

DOI: 10.1177/014616702237644

© 2002 by the Society for Personality and Social Psychology, Inc.

male counterparts, still avoid majors involving moderate to high levels of math (Lefevre, Kulak, & Heymans, 1992). Lefevre and her colleagues (1992) found that among these equally gifted students, only 19% of women reported choosing majors that involved even a moderate level of mathematics (e.g., architecture, business, and economics), whereas 43% of men reported choosing majors in these domains. This gender divide was magnified when the researchers focused on majors considered high in math content (e.g., computer science, engineering, and mathematics), in which men are 4 1/2 times more likely than women to pursue majors. Exacerbating this problem, the college drop-out rate for women in mathematics, engineering, and the physical sciences is 2 1/2 times that of men (Hewitt & Seymour, 1991). Naturally, these gender differences extend beyond the classroom to the workplace, where disproportionate numbers of men and women enter careers in quantitative fields (Eccles, 1994; Stangor & Sechrist, 1998). Even though there has been a major increase in the participation of women in the general workforce, some traditionally masculine domains (e.g., engineering and physics) have seen no significant increase in women's participation over the past 50 years (Culotta & Gibbons, 1992). These gender differences in career selection are most pronounced in the fields of math, engineering, and the physical sciences, in which women occupy only 10% of the jobs and earn just 75% of what their male counterparts earn (Crocker, Major, & Steele, 1998; Hewitt & Seymour, 1991).

Women in traditionally masculine fields must contend with the added burden that accompanies cultural stereotypes alleging a sex-based inability. Cultural stereotypes are prevailing attitudes about the attributes of stigmatized groups that are conveyed through societal-level messages and images, such as those broadcast by our mass media (Ashmore, Del Boca, & Wohlers, 1986). Despite widespread concern in our society over the consequences of being raised on a steady diet of cultural stereotypes, there has been remarkably little experimental research conducted on the effects of exposure to gender-stereotypic media images. Numerous correlational studies, however, have documented the consequences of heavy television consumption. For example, Gerbner, Gross, Morgan, and Signorielli (1993) found that individuals who watch a great deal of television, compared to light viewers, believe that women have limited abilities and interests compared to men. These researchers also found that heavy viewers expressed more stereotypic views regarding what professions are best suited for women. Even more incriminating, a naturalistic experiment found that after television was introduced to a small community, the children's attitudes became more

gender stereotypic, mirroring those of children raised on television (Kimball, 1986).

STEREOTYPE THREAT

Negative stereotypes are so pervasive in our mass-media culture that those who are stigmatized have detailed knowledge of the accusations targeted at their group.

African Americans, for example, are likely to be well aware that stereotypes accuse them of being intellectually inferior and aggressive; and women are well aware that stereotypes accuse them of being emotional, bad at math, and lacking leadership aptitude. (Crocker et al., 1998, p. 518)

This detailed knowledge of relevant stereotypes can lead to a situational predicament for stigmatized individuals (Steele, 1997). Specifically, the risk of being personally reduced to a negative stereotype can elicit a disruptive state among stigmatized individuals that undermines performance and aspirations in any alleged stereotype-relevant domain—a situational predicament termed “stereotype threat” (Spencer, Steele, & Quinn, 1999; Steele, 1997; Steele & Aronson, 1995). When members of a stigmatized group find themselves in a situation where negative stereotypes provide a plausible explanation for their behavior, stereotype threat is the unsettling awareness that they risk being judged by, or treated in terms of, those negative stereotypes (Steele, 1997). Whereas susceptibility to stereotype threat requires the stigmatized to have specific knowledge of negative stereotypes and the domains to which they apply, this situational predicament does not require the stigmatized to have any internal doubts about their ability, or their group's ability, in those domains. In fact, the effects of stereotype threat may be most acutely felt by those individuals who are invested and skilled in the targeted domain, or by those individuals who at least care about the social consequences of being judged incompetent in that domain (Aronson, Quinn, & Spencer, 1998; Steele, 1997).

If this situational predicament is the root of underachievement in targeted domains, then eliminating the situational factors that give rise to stereotype threat should enable stigmatized individuals to perform to their full potential. Steele and Aronson (1995) tested this notion by varying the degree of stereotype threat that African American students experienced during a difficult verbal test. As predicted, when participants were told the test was nondiagnostic of intellectual ability, which makes the intellectual-inferiority stereotype irrelevant to the exam, Black and White students performed equally well on the test. When participants were told the

test was diagnostic of intellectual ability, however, the traditional gap in performance between Black and White students resurfaced. The researchers then examined whether priming African American students to think about their race could expose them to the detrimental effects of stereotype threat even on a nondiagnostic test. Steele and Aronson (1995) confirmed that making race salient, which elicits the threat of being judged in terms of racial stereotypes, undermined the African American students' performance even on the nondiagnostic version of the test.

Because cultural stereotypes question women's ability to succeed in any traditionally masculine domain, women are susceptible to stereotype threat in all fields considered inherently masculine. Again, if this situational predicament is the basis for women's performance deficits, then removing stereotype threat from traditionally masculine domains should eliminate women's underachievement. Spencer et al. (1999) tested this hypothesis by manipulating the level of stereotype threat that their female participants experienced while taking a difficult math test. Half of the participants were told that the math test had revealed "no-gender-differences" in the past, which makes the math-inferiority stereotype irrelevant to the exam, whereas participants in the control condition were provided with no information about the relative past performance of men and women. Not surprisingly, the well-documented gender gap in math performance emerged among participants in the control condition. When stereotype threat was eliminated from the testing situation via the no-gender-difference instructions, however, women and men performed equally well on the difficult math test.

MEDIATORS OF STEREOTYPE-THREAT EFFECTS

Stereotype threat is a multifaceted situational predicament whose mediational path can be shaped by features of the person, the context, and their interaction (see Steele, Spencer, & Aronson, in press). Despite these complexities, considerable progress has been made investigating potential mediators of stereotype-threat effects (e.g., effort, expectancies, biological processes, stereotype activation, evaluation apprehension, anxiety, etc.). Of these potential mediators, anxiety has arguably received the most attention, but reliable evidence for the mediational role of anxiety has been elusive (e.g., Leyens, Désert, Croizet, & Darcis, 2000; Osborne, 2001; Spencer et al., 1999; Stangor, Carr, & Kiang, 1998; Stone, Lynch, Sjomeling, & Darley, 1999). Of course, anxiety is only a potential mediator of stereotype-threat effects when the target is actually experiencing, or at least contemplating, a threatening domain. Researchers overlooking this fact may have underestimated the reliability

of anxiety as a mediator. Relying on self-reported measures of anxiety also may have contributed to the inconsistency of previous mediational findings—participants may not always be willing or able to accurately report their own level of anxiety. Blascovich, Spencer, Quinn, and Steele (2001) tested this notion by collecting both self-report and physiological indicators of anxiety (i.e., blood pressure) while Black and White participants took a difficult verbal test described as being either culturally unbiased or diagnostic of ability. Replicating previous stereotype-threat findings, the African American participants only underperformed on the diagnostic test. According to the self-report measures, their under-performance was not accompanied by an increase in anxiety; however, the physiological measures confirmed that their underperformance was actually accompanied by a significant increase in blood pressure. Therefore, both the timing and sensitivity of measures will be critical in determining whether anxiety is a reliable mediator of stereotype-threat effects.

The present research will test another potential mediator of stereotype-threat effects, one which Steele and Aronson (1995) exposed when they discovered that African Americans facing a diagnostic test activated self-relevant stereotypes. Steele and Aronson did not administer the above test; therefore, a full mediational analysis was not possible. Nevertheless, we believe that susceptibility to stereotype threat involves confronting activated self-relevant stereotypes and knowing that one risks being personally reduced to those stereotypes in targeted domains. Thus, activation level of self-relevant stereotypes should reliably mediate the effects of stereotype threat. Testing this particular mediational pathway also allows us to clarify the differences between the stereotype-threat paradigm and the ideomotor paradigm. Ideomotor theory suggests that primed stereotypes can automatically evoke stereotype-consistent behavior, unbeknownst to the individual displaying that behavior (e.g., Bargh, Chen, & Burrows, 1996; Dijksterhuis et al., 1998). For example, Bargh and his colleagues (1996) showed that undergraduates subtly primed with an elderly stereotype walked more slowly to an elevator after leaving the laboratory. According to the ideomotor perspective, because behavior stems mechanically and unconsciously from primed stereotypes, whether stereotypes are self-relevant to the target is immaterial—all targets cognizant of the stereotypes are equally susceptible to ideomotor effects (see Wheeler & Petty, 2001). In contrast, our stereotype-threat paradigm requires targets to confront activated self-relevant stereotypes and to realize that they could be personally reduced to those stereotypes in a given situation. Therefore, stereotype relevance should moderate whether level of stereotype activation mediates our stereotype-threat effects.

THE PRESENT RESEARCH

In Study 1, we will explicitly prime the female stereotype among both men and women. Because stereotype threat partially stems from the broad dissemination of cultural stereotypes in our society, we decided to prime the female stereotype through exposure to gender-stereotypic television commercials. After viewing the commercials, we predict that only participants for whom the activated stereotype is self-relevant (i.e., women) will underperform on a subsequent math test and that level of stereotype activation among those women will mediate the performance-inhibiting effects of the commercials.

The assumption underlying Studies 2 and 3 is that women who risk being personally reduced to a negative stereotype may withdraw from targeted domains in an attempt to cope with the disruptive state they can elicit. This defensive detachment may not only undermine women's short-term performance in traditionally masculine domains but also undermine their long-term aspirations and accomplishments in those domains (Crocker et al., 1998; Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Steele, 1997). Specifically, we believe that stereotype threat can lead women to withdraw from quantitative domains while seeking domains in which they do not risk being negatively stereotyped. In Study 2, we predict that exposure to the gender-stereotypic commercials prior to a nondiagnostic aptitude test will induce women to avoid math items in favor of verbal items. As Steele (1997) succinctly put it, "Women may reduce their stereotype threat substantially by moving across the hall from math to English class" (p. 618). In Study 3, women exposed to the gender-stereotypic commercials are expected to subsequently indicate less interest in pursuing those educational and vocational options in which they are susceptible to stereotype threat (i.e., quantitative domains) while indicating more interest in those domains that do not allege a sex-based inability (i.e., verbal domains).

STUDY 1

As discussed earlier, Steele and Aronson (1995) discovered that informing African American participants that a test was nondiagnostic of intellectual ability eliminated stereotype threat from the situation. Prior to running Study 1, a pilot study was required to confirm that similar nondiagnostic instructions could eliminate women's underperformance on a difficult math test.

PILOT STUDY

Eighty-three undergraduates were randomly assigned to a 2 (gender) \times 2 (instruction type) factorial design, with the dependent variable being performance on a difficult math test—displayed as percentage correct. Partic-

ipants read instructions that either stated that the math test was nondiagnostic of ability or provided no diagnostic information. A 2 (gender) \times 2 (instruction type) ANOVA on the performance data revealed a main effect for gender, $F(1, 79) = 5.34, p < .05$, and a main effect for instruction type, $F(1, 79) = 4.07, p < .05$, but these main effects were qualified by a significant two-way interaction between Gender and Instruction Type, $F(1, 79) = 5.57, p < .05$. Among participants who read the nondiagnostic instructions, simple-effect tests revealed that women and men performed equally well on the difficult math test ($M = 25\%$ and $M = 25\%$, respectively), $F < 1$. When provided with no diagnostic information, however, the traditional performance gap reappeared between women and men ($M = 12\%$ and $M = 26\%$, respectively), $F(1, 79) = 14.41, p < .01$. These results, which are a conceptual replication of Steele and Aronson's (1995) findings, verify that nondiagnostic instructions can eliminate women's underperformance on difficult math tests.

For Study 1, before taking a difficult nondiagnostic math test, half of our participants were exposed to gender-stereotypic television commercials designed to explicitly prime the female stereotype. We predicted that only participants for whom the activated stereotype was self-relevant (i.e., women) would underperform on the subsequent math test. Furthermore, we expected that level of stereotype activation among those women would mediate the performance-inhibiting effects of the commercials. To test this mediational hypothesis, activation of the female stereotype was measured following exposure to the commercials but prior to the math test. This design allowed us to test whether exposure to the commercials activated the female stereotype and whether the resulting level of stereotype activation mediated the effects of those commercials.

Method

PARTICIPANTS AND DESIGN

Participants were 54 undergraduates (29 men, 25 women) at the University of Waterloo who were enrolled in a second semester calculus course. All participants had completed just one semester of calculus and had received a grade of "B" or better in that initial course. Selected participants also indicated that they strongly agreed with both of the following statements: (a) I am good at math and (b) It is important to me that I am good at math. As mentioned earlier, the effects of stereotype threat may be most acutely felt by those individuals who are skilled in the domain or at least care about the stigma of being labeled incompetent in that domain (Aronson et al., 1999; Steele, 1997). Approximately equal numbers of men and women were randomly assigned to the 2 (gender) \times 2 (commercial type) factorial design. The primary dependent variables were per-

formance on the math test and activation of the female stereotype as measured by a lexical-decision task.

MATERIALS

Television commercials. Participants were exposed to either a stereotypic or counterstereotypic set of television commercials prior to completing a lexical-decision task and then a nondiagnostic math test. All of the advertisements employed in this research were actual television commercials that aired during major network broadcasts. Both sets of commercials contained six ads, four of which were neutral. The same four neutral commercials appeared in both sets and did not advertise any gender-stereotypic products or companies; no humans were depicted in those commercials. The products and companies advertised in the neutral commercials included a cellular phone, a gas station, a pharmacy, and finally an insurance company. One of the gender-stereotypic commercials portrayed a young woman who was so excited about being a consumer of a new acne product that she bounced on her bed with joy, whereas the other portrayed a woman "drooling" with anticipation to try a new brownie mix. It should be noted that neither of the gender-stereotypic commercials made any reference to alleged gender differences in math ability or major/career selection. The first counterstereotypic commercial portrayed an attractive woman impressing a man with her knowledge of automotive engineering and the second revealed a woman speaking intelligently about health care concerns.

Assessing stereotype activation. The lexical-decision task requires participants to identify as quickly as possible whether letter strings flashed on a computer screen are words or nonwords. The premise behind the lexical-decision measure is that participants for whom the female stereotype is activated should be quicker to recognize words related to that stereotype than participants for whom that stereotype is not activated. The task was conducted on Macintosh LC series computers programmed with Superlab software. The stimuli were presented in the center of the screen as black words on a white background. Each letter string remained on the screen until the participant responded. The letter strings used for this lexical-decision task were words related to the female stereotype, neutral words, and nonwords. The order of presentation for these letter strings was randomized for each participant, and just more than half of all these letter strings were nonwords.

The stereotypic words employed in this study were generated during pretesting by students enrolled in an undergraduate psychology course at the University of Waterloo. The male and female undergraduates were instructed to generate a list of adjectives to characterize the female stereotype. The final 10 stereotypic words

employed in this study were *intuitive, gullible, irrational, wasteful, inferior, distracted, emotional, indecisive, tense, and weak*. Each of these stereotypic words was then matched with a neutral word of similar length and frequency based on norms reported by Kucera and Francis (1967). The purpose of matching the stereotypic words with neutral words is to control for individual differences in reaction time, which is accomplished by using these matched neutral words as a covariate when analyzing the composite of stereotypic words (Kunda, Davies, Adams, & Spencer, 2002).

Math test. The 12 questions on the math test were taken from Advanced Subject Graduate Record Examinations (GREs) in Mathematics (Educational Testing Service, 1987). These particular items were selected to ensure that the participants would find the test sufficiently difficult. The math test was administered on the same Macintosh LC series computers employed for the lexical-decision task. For each math question, participants had the opportunity to answer the question, leave the question blank, or skip the question, allowing them to answer it later. The standard Educational Testing Service (ETS) formula was employed for scoring the math test; correct items receive 1 point, incorrect items receive 1/5-point deduction, and items left blank receive neither points nor deductions.

PROCEDURE

Prior to the participants reporting to the laboratory, the experimenter randomly chose one of two unmarked videos for the upcoming session. One video contained the stereotypic set of commercials and the other contained the counterstereotypic set of commercials. The experimenter ran the participants in mixed-gender groups ranging in size from two to five people. As the participants arrived, they were informed that the experimenter was interested in testing their long-term memory of details contained in television commercials. Ostensibly to achieve this goal, the experimenter asked the participants to watch a short video consisting of several commercials. The experimenter then turned on the television, cued the previously selected video on the VCR, and left the room. The first 10 s of both videos revealed only a blue screen, providing ample time for the experimenter to vacate the room prior to the start of the commercials. When the experimenter returned to the laboratory after the allotted 3 min, the commercials were finished and both videos again revealed only a blue screen. This design enabled the experimenter to remain blind to the participants' condition throughout the study.

After returning to the laboratory, the experimenter informed the participants that approximately 30 min had to elapse before their long-term memory of the com-

mercials could be accurately tested. To fill this time before the memory test, the participants were asked to complete two cognitive-busy tasks on the Macintosh LC series computers contained in the room. The first cognitive-busy task was actually the lexical-decision task designed to measure activation of the female stereotype, which only took 5 min to complete. Instructions prior to the lexical-decision task informed participants to respond as quickly and accurately as possible. The second cognitive-busy task was the math test. Instructions prior to the math test informed participants that the current version of the test was nondiagnostic of ability. Participants also were informed that they had 20 min to complete as many of the math problems as possible. After the allotted 20 min, the experimenter confirmed that all participants still believed that their memory for the details contained in the television commercials was being tested. The experimenter then informed the participants that the study was over. The subsequent debriefing session contained two manipulation checks: (a) The experimenter confirmed that all participants realized they had been exposed to either stereotypic or counterstereotypic portrayals of women in their assigned commercial set and (b) the experimenter confirmed that all participants knew of the math-inability stereotype targeted at women.

Results and Discussion

Participants were exposed to either stereotypic or counterstereotypic commercials prior to completing a lexical-decision task and then a nondiagnostic math test. Exposure to the stereotypic commercials was expected to prime the female stereotype among both men and women. Only women, however, for whom the activated stereotype was self-relevant were expected to underperform on the subsequent math test. Furthermore, level of stereotype activation among the female participants was expected to mediate the performance-inhibiting effects of those commercials. To test these predictions, two separate dependent variables were required: (a) math performance, calculated using the standard ETS scoring formula, and (b) activation of the female stereotype, as measured by a lexical-decision task.

MATH PERFORMANCE

We predicted that only women exposed to the gender-stereotypic commercials would underperform relative to men on the nondiagnostic math test. A 2 (gender) \times 2 (commercial type) ANOVA on the performance data revealed a main effect for gender, $F(1, 50) = 6.91, p < .05$, but this main effect was qualified by a significant two-way interaction between Gender and Commercial Type, $F(1, 50) = 4.38, p < .05$. As expected, simple-effect tests on the performance data revealed that women and men per-

formed equally well ($M = 31%$ and $M = 34%$, respectively) on the difficult math test following exposure to the counterstereotypic ads, $F < 1$. When exposed to the stereotypic ads, however, women performed significantly worse than men on the nondiagnostic test ($M = 19%$ and $M = 39%$, respectively), $F(1, 50) = 11.57, p < .01$. Further simple-effect tests confirmed that women in the stereotypic condition underperformed on the math test compared to women in the counterstereotypic condition, $F(1, 50) = 4.71, p < .05$.

STEREOTYPE ACTIVATION

As measured by their recognition times for female stereotypic words on the lexical-decision task, it was predicted that both male and female participants who viewed the gender-stereotypic commercials would exhibit activation of the female stereotype compared to those participants who viewed the counterstereotypic commercials. Trials in which the participant responded incorrectly, or after the 2,000-ms time limit, were excluded from analyses (4.4% of trials). In addition, trials in which the response time exceeded 2.36 standard deviations from the trial's mean response time (3.9% of trials) also were considered outliers and excluded from analyses (Van Selst & Jolicoeur, 1994). There were no differences between conditions in number of errors or outliers, all $F_s < 1$.

A 2 (gender) \times 2 (commercial type) ANCOVA was conducted on the participants' recognition time for female stereotypic words, with the covariate being recognition time for matched neutral words (Kunda et al., 2002). The ANCOVA revealed only the predicted main effect for commercial type, $F(1, 49) = 11.45, p = .001$. Regardless of gender, participants who were exposed to the stereotypic commercials activated the female stereotype compared to those participants exposed to the counterstereotypic commercials. Simple-effect tests confirmed there were no differences in activation levels between men and women in the stereotypic condition, $F < 1$, or the counterstereotypic condition, $F < 1$. Additional simple-effect tests revealed that recognition times for female stereotypic words were faster among women in the stereotypic condition ($M = 701$ ms) than women in the counterstereotypic condition ($M = 784$ ms), $F(1, 49) = 6.27, p < .05$. Similarly, recognition times for female stereotypic words were faster among men in the stereotypic condition ($M = 718$ ms) than men in the counterstereotypic condition ($M = 778$ ms), $F(1, 49) = 3.83, p = .05$.

MEDIATION

To test whether level of stereotype activation mediated the effect of the commercials on math performance, we conducted a series of regression analyses recommended by Baron and Kenny (1986) and Judd and

Kenny (1981). Because our stereotype-threat manipulation did not affect men's math performance, these regression analyses were conducted solely on the women's data. Once again, when analyzing the composite of stereotypic words, individual differences in reaction time are controlled for by using matched neutral words as a covariate (Kunda et al., 2002).

There are four steps necessary to test whether level of stereotype activation among women mediated the effect of commercial type on math performance. First, we conducted a regression analysis to determine if commercial type actually predicted math performance, which was confirmed, $\beta = .41$, $t_{(24)} = 2.17$, $p < .05$. Next, we examined whether commercial type predicted activation of the female stereotype, which also was confirmed, $\beta = .54$, $t_{(24)} = 3.07$, $p < .01$. A third step examined whether stereotype activation (the potential mediator) predicted math performance when controlling for the effect of commercial type. Results confirmed that level of stereotype activation did in fact predict performance, $\beta = .48$, $t_{(23)} = 2.30$, $p < .05$. The final, and most crucial, step establishes whether stereotype activation actually mediated the effect of commercial type on math performance. The amount of mediation is defined as the amount of reduction in the effect of commercial type on math performance when controlling for stereotype activation. When level of stereotype activation was controlled for in the analysis, the effect of commercial type on math performance dropped from $\beta = .41$, $t_{(24)} = 2.17$ to $\beta = .16$, $t_{(23)} < 1$. Employing the Goodman (1960) equation verified that this degree of mediation was indeed significant, $Z = 1.93$, $p = .05$. Taken together, the above analyses confirmed that level of stereotype activation among women did mediate the effect of the commercials on math performance.

Study 1 revealed that exposure to the gender-stereotypic commercials resulted in the activation of the female stereotype among both men and women; however, only women for whom the activated stereotype was self-relevant underperformed on the subsequent nondiagnostic math test. Furthermore, it was discovered that level of stereotype activation among the female participants mediated the detrimental effects of the television commercials. This study confirmed that susceptibility to stereotype threat is limited to those individuals confronting activated self-relevant stereotypes.

To ensure that taking the lexical-decision task did not augment the effects of viewing the commercials, we decided to replicate Study 1 without the lexical-decision task. Seventy-five undergraduates were randomly assigned to a 2 (gender) \times 2 (commercial type) factorial design, with the dependent variable being performance on a nondiagnostic math test. A 2 (gender) \times 2 (commercial type) ANOVA on the performance data revealed a

main effect for gender, $F(1, 71) = 3.87$, $p < .05$, but this main effect was qualified by a significant two-way interaction between Gender and Commercial Type, $F(1, 71) = 7.78$, $p < .01$. As expected, simple-effect tests on the performance data revealed that women and men performed equally well ($M = 20.8\%$ and $M = 18.2\%$, respectively) on the difficult math test following exposure to the counterstereotypic ads, $F < 1$. When exposed to the stereotypic ads, however, women again underperformed compared to men ($M = 10.9\%$ and $M = 25.9\%$, respectively), $F(1, 71) = 8.33$, $p < .01$. These findings replicate our earlier research and suggest that the act of taking the lexical-decision task did not influence our results in Study 1.

STUDY 2

The goal of Study 2 was to explore whether stereotype threat could lead women to avoid the domain of math while seeking a domain in which they are immune to stereotype threat. Because cultural stereotypes do not accuse women of having inferior verbal skills, women in verbal domains do not risk being personally reduced to a negative stereotype. Therefore, we decided to examine whether stereotype threat could persuade women to leave quantitative domains in favor of verbal domains. Study 1 confirmed that exposure to gender-stereotypic commercials, which prime the female stereotype, evokes stereotype threat among women taking a nondiagnostic math test. Consequently, we hypothesized that exposure to gender-stereotypic commercials would induce women taking a nondiagnostic aptitude test to avoid math items in favor of verbal items. Therefore, the math test employed for Study 1 was replaced with an aptitude test for Study 2. Furthermore, because the counterstereotypic-commercial condition in Study 1 simply replicated our previous no-commercial findings (see Pilot Study data), for Study 2 we replaced the counterstereotypic commercials with neutral commercials.

Method

PARTICIPANTS AND DESIGN

Participants were 67 undergraduates (33 men, 34 women) at the University of Waterloo who participated for credit in an introductory psychology class. On a mass testing questionnaire administered earlier in the term, selected participants indicated that they strongly agreed with both of the following statements: (a) I am good at math and (b) It is important to me that I am good at math. Approximately equal numbers of men and women were randomly assigned to the 2 \times 2 \times 2 mixed-model design, which had two between-participants factors (gender and commercial type) and one within-participants factor (question type). The primary dependent

variable was the relative number of verbal and math items attempted, and the secondary dependent variable was relative performance on the verbal and math portions of the aptitude test.

MATERIALS

The 24 verbal and 24 math questions contained in the aptitude test were compiled from several previous versions of the GRE General Test (Educational Testing Service, 1992). Only questions with an ETS correct response rate ranging from 20% to 70% were selected for the aptitude test, which guaranteed that the participants would find the test sufficiently difficult. To ensure that the verbal questions were no more appealing than the math questions, the verbal portion of the exam was constructed to be slightly more difficult than the math portion. The mean ETS correct response rate was 47% for the verbal portion and 52% for the math portion. The test was composed of four verbal sections and four math sections, each consisting of six questions. The math and verbal sections were alternated, and which came first was counterbalanced.

The counterstereotypic commercials employed in Study 1 were removed for Study 2; therefore, participants either viewed the stereotypic-commercial set or a new neutral-commercial set. The four commercials contained in the neutral-commercial set, however, are the same four neutral commercials employed in Study 1. The set of advertisements viewed in the stereotypic-commercial condition again included those four neutral commercials intermixed with two gender-stereotypic commercials.

PROCEDURE

The procedures for Study 1 and Study 2 were identical up to the point at which the experimenter (blind to condition) returned to the laboratory after the participants had viewed their randomly assigned set of commercials. The participants, who were run in mixed-gender groups ranging in size from two to five people, were informed by the experimenter that 20 min had to elapse before their long-term memory for the commercials could be accurately tested. The experimenter told the participants that rather than having them waste their time on a filler task, we would appreciate them volunteering to participate in the creation of a new standardized test being developed at the University of Waterloo. They were informed that this research was just beginning and the developers were trying to determine which questions should be included in future versions of the test. The participants also were explicitly told that the current version of the test was nondiagnostic of ability.

All participants volunteered to take the test. Prior to starting, the participants were instructed to carefully read the test's coversheet, which explained the composi-

tion of the test and informed the participants that they had 20 min to attempt as many questions as possible. The coversheet also indicated that the standard ETS scoring formula would be employed for the test; correct items receive 1 point, incorrect items receive 1/5-point deduction, and items left blank receive neither points nor deductions. The final instructions read by the participants were as follows:

To accurately evaluate this problem set we need to know which questions you worked on and which questions you skipped. So if you work on a problem and decide not to answer it, please put a checkmark next to that item. Having a record of the items not attempted by students (i.e., no checkmark or answer) will provide us with extremely helpful information when determining which questions should be included in future versions of this test.

After the participants had worked on the test for the allocated 20 min, the experimenter confirmed that all participants still believed their memory for the details contained in the television commercials was being tested. The subsequent debriefing session contained two manipulation checks: (a) The experimenter confirmed that participants in the stereotypic condition realized they had been exposed to stereotypic portrayals of women in their assigned commercial set and (b) the experimenter confirmed that participants were aware of the math-inferiority stereotype targeted at women.

Results and Discussion

The relative number of math and verbal items attempted by women in the neutral-commercial condition was expected to be similar to that of men in either condition. In contrast, women exposed to the gender-stereotypic commercials were expected to avoid the math items in favor of the verbal items. Second, we also expected that women in the stereotypic-commercial condition would replicate our Study 1 findings by underperforming on the math portion of the aptitude test. To test these predictions two separate dependent variables were required: (a) the relative number of math and verbal items attempted (i.e., checkmark or answer) and (b) the performance score calculated using the standard ETS scoring formula.

ATTEMPTS

We hypothesized that our math-invested participants would naturally attempt more math items than verbal items on the nondiagnostic aptitude test. In comparison, women exposed to the gender-stereotypic commercials were expected to avoid the math questions in favor of the verbal questions. A three-way ANOVA (Gender \times Commercial Type \times Question Type) on the number of items attempted revealed the predicted main effect for ques-

tion type, $F(1, 63) = 4.40, p < .05$, revealing the participants' natural tendency to attempt more math than verbal questions on an aptitude test. That main effect, however, was qualified by a marginal two-way interaction between Commercial Type and Question Type, $F(1, 63) = 3.65, p = .057$, which in turn was qualified by a significant three-way interaction between Gender, Commercial Type, and Question Type, $F(1, 63) = 6.43, p = .01$.

We expected men's behavior not to be influenced by our commercial-type manipulation; therefore, we broke down the above triple interaction by gender and separately analyzed the data for our male and female participants. Examining the men's data, a two-way ANOVA (Commercial Type \times Question Type) on the number of items attempted revealed only the predicted main effect for question type, $F(1, 63) = 6.63, p = .01$. That is, regardless of condition, men attempted significantly more math questions ($M = 15.19$) than verbal questions ($M = 12.33$). In contrast, a two-way ANOVA (Commercial Type \times Question Type) on the number of items attempted by women revealed only the predicted two-way interaction between Commercial Type and Question Type, $F(1, 63) = 10.04, p < .01$.

Simple-effect tests on the women's data revealed that women exposed to neutral commercials, similar to men in either condition, attempted significantly more math than verbal questions ($M = 15.47$ and $M = 11.59$, respectively), $F(1, 63) = 6.28, p = .01$, whereas women exposed to gender-stereotypic commercials revealed the exact opposite pattern of results, attempting significantly more verbal than math questions ($M = 15.29$ and $M = 12.24$, respectively), $F(1, 63) = 3.90, p < .05$. Additional simple-effect tests indicated that women in the stereotypic-commercial condition attempted marginally fewer math questions, $F(1, 63) = 3.01, p = .08$, but more verbal questions, $F(1, 63) = 3.95, p < .05$, than women in the neutral-commercial condition (see Figure 1).

PERFORMANCE

Because all of the participants selected for this study were highly invested in the domain of math, we expected their math performance on the nondiagnostic aptitude test would naturally be stronger than their verbal performance. Based on our findings in Study 1, however, we predicted that women in the stereotypic-commercial condition would underperform on the math portion of the test. A three-way ANOVA (Gender \times Commercial Type \times Question Type) on the performance data revealed a main effect for question type, $F(1, 63) = 61.98, p < .001$. This main effect verified that the participants' math performance on average was indeed stronger than their verbal performance. However, this main effect was qualified by a significant two-way interaction between Gender and Question Type, $F(1, 63) = 8.11, p < .01$, which

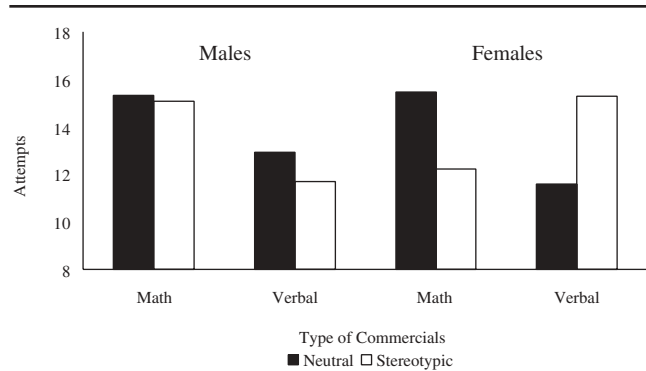


Figure 1 Math and verbal attempts as a function of gender and commercial type.

in turn was qualified by a significant three-way interaction between Gender, Commercial Type, and Question Type, $F(1, 63) = 6.91, p = .01$.

Once again, we broke down this triple interaction by gender and separately analyzed the performance data for our male and female participants. A two-way ANOVA (Commercial Type \times Question Type) on the men's performance data revealed only the predicted main effect for question type, $F(1, 63) = 56.59, p < .001$. That is, regardless of condition, men did significantly better on the math portion ($M = 39.38\%$) of the aptitude test than on the verbal portion ($M = 11.57\%$). Turning to the women's data, a two-way ANOVA (Commercial Type \times Question Type) on the performance data also revealed a main effect for question type, $F(1, 63) = 12.82, p < .001$, but that main effect was qualified by a significant two-way interaction between Commercial Type and Question Type, $F(1, 63) = 9.09, p < .01$.

Simple-effect tests on the women's performance data revealed that women exposed to the neutral commercials, similar to men in either condition, did significantly better on the math portion of the exam than the verbal portion ($M = 34.80\%$ and $M = 10.83\%$, respectively), $F(1, 63) = 21.75, p < .001$. Exposure to the gender-stereotypic commercials, however, completely eliminated this math-over-verbal performance bias ($M = 20.70\%$ and $M = 18.62\%$, respectively), $F < 1$. Further simple-effect tests confirmed that women in the stereotypic-commercial condition underperformed on the math portion of the test compared to women in the neutral-commercial condition, $F(1, 63) = 7.02, p < .01$. It should be noted that these performance results are not simply a by-product of the number of items attempted. In fact, when controlling for number of attempts, a very similar pattern of data emerges, including the finding that exposure to gender-stereotypic commercials undermines women's math performance.¹

Taken together, these findings provide strong initial evidence that stereotype threat can lead women to temporarily avoid traditionally masculine domains and seek domains where they do not risk being personally reduced to a negative stereotype. Specifically, it was shown that exposure to gender-stereotypic television commercials led women to avoid math items in favor of verbal items on a subsequent nondiagnostic aptitude test. Replicating Study 1, women exposed to these gender-stereotypic commercials also underperformed on the math portion of that nondiagnostic test. Whereas cultural stereotypes accuse women of having inferior math skills, no such stereotypes allege inferior verbal skills; therefore, women are not vulnerable to stereotype threat in verbal domains.

STUDY 3

The primary objective of Study 3 was to expand our paradigm to test whether stereotype threat also could influence women's educational and vocational aspirations. The assumption being that stereotype threat would lead women to temporarily withdraw from traditionally masculine domains in favor of domains in which stereotypes do not allege a sex-based inability. To examine this issue, we replaced the aptitude test employed in Study 2 with a survey in which participants indicated their current interest in a range of college majors and careers. Women exposed to the gender-stereotypic commercials were expected to subsequently indicate less interest in pursuing educational/vocational options in which they would likely experience stereotype threat (i.e., quantitative domains) and more interest in those options in which they are immune to stereotype threat (i.e., verbal domains).

Method

PARTICIPANTS AND DESIGN

Participants were 62 undergraduates (30 men, 32 women) at the University of Waterloo who participated for credit in an introductory psychology class. The participants were drawn from the same subject pool utilized in Study 2. Approximately equal numbers of men and women were randomly assigned to the $2 \times 2 \times 2$ mixed-model design, which had two between-participants factors (Gender and Commercial Type) and one within-participants factor (Domain Type). The primary dependent variable was current degree of interest indicated for majors/careers in quantitative domains relative to verbal domains.

MATERIALS

The identical sets of neutral and gender-stereotypic commercials used in Study 2 were again employed in Study 3. After watching their assigned set of commer-

cial, the participants completed a survey. The survey asked participants to indicate their current degree of interest in dozens of educational and vocational alternatives. Participants indicated their preferences using a scale ranging from 1 (*no interest*) to 7 (*strong interest*). Apart from the filler items, the majors and careers selected for the survey were intended to represent a cross section from quantitative and verbal domains. For example, alternatives drawn from the fields of mathematics, engineering, and computer science were countered with alternatives drawn from the fields of creative writing, communications, and linguistics.

PROCEDURE

The procedures for Study 3 were identical to Studies 1 and 2 up to the point at which the experimenter (blind to condition) returned to the laboratory after the participants had viewed their randomly assigned set of commercials. The participants, who were run in mixed-gender groups ranging in size from two to five people, were informed by the experimenter that 20 min had to elapse before their long-term memory for the commercials could be accurately tested. Supposedly to fill those 20 min, the students were asked to voluntarily participate in an important survey ostensibly being conducted by Career Services at the University of Waterloo. The survey's coversheet claimed that the results of the campus-wide survey would directly affect students by determining future services and programs to be offered by the university. All participants volunteered to complete the survey. Once they had completed the survey, the experimenter confirmed that all participants still believed their memory for the details contained in the television commercials would be tested. The subsequent debriefing session contained two manipulation checks: (a) The researcher confirmed that participants in the stereotypic condition realized they had been exposed to stereotypic portrayals of women and (b) the researcher confirmed that participants knew of the stereotype that women have poor quantitative skills.

Results and Discussion

With dozens of educational and vocational alternatives listed on the survey, we decided to employ factor analysis to help us identify a relatively small number of meaningful factors within that data. To allow for correlations among the underlying factors, we conducted a factor analysis with an oblimin-rotated solution on the preference data for 57 educational/vocational alternatives. The scree plot of the resulting eigenvalues indicated that a three-factor solution would be most appropriate for the preference data. Two of the three resulting factors represent polar opposite domains with respect to stereotype-threat susceptibility. The Quantitative

Domain encompasses 25 educational/vocational alternatives (Cronbach's $\alpha = .91$) that require mathematical skills to succeed (e.g., engineer, mathematician, computer science, statistics, accountant, physics, etc.). The Verbal Domain is composed of 20 educational/vocational alternatives (Cronbach's $\alpha = .92$) that rely more heavily on verbal skills for success (e.g., author of novels, linguistics, journalist, communications, political science, editor, etc.). The third factor to emerge, the Health Domain, consists of 12 educational/vocational alternatives (Cronbach's $\alpha = .92$) in the health sciences (e.g., nurse, pharmacist, optometrist, physician, kinesiology, etc.). All three factors in their entirety are reported in Davies (2000). Because our primary interest in this study was to determine whether exposure to gender-stereotypic commercials could undermine women's aspirations in traditionally masculine fields while encouraging them to pursue fields in which they are immune to stereotype threat, we will limit our analyses to the Quantitative and Verbal Domains.²

PREFERENCES

We expected women in the neutral-commercial condition to reveal a similar pattern of interest as men in either condition, that is, a preference for the Quantitative Domain, which encompasses the most popular majors at the University of Waterloo. Women in the stereotypic-commercial condition, however, were expected to reveal less interest in the Quantitative Domain and more interest in the Verbal Domain. A three-way ANOVA (Gender \times Commercial Type \times Domain Type) on the preference data revealed a significant two-way interaction between Gender and Domain Type, $F(1, 58) = 22.05, p < .001$, and a significant two-way interaction between Commercial Type and Domain Type, $F(1, 58) = 17.98, p < .001$, which were both qualified by a significant three-way interaction between Gender, Commercial Type, and Domain Type, $F(1, 58) = 4.05, p < .05$.

Because we expected men's preferences not to be influenced by our commercial-type manipulation, we broke down the above triple interaction by gender, separately analyzing the data for our male and female participants. A two-way ANOVA (Commercial Type \times Domain Type) on the men's preference data revealed only the predicted main effect for domain type, $F(1, 58) = 11.58, p < .01$. Regardless of condition, men expressed more interest in the Quantitative Domain ($M = 3.19$) than the Verbal Domain ($M = 2.48$). Turning to the women's data, a two-way ANOVA (Commercial Type \times Domain Type) also revealed a main effect for domain type, $F(1, 58) = 10.35, p < .01$, but this main effect was qualified by a significant two-way interaction between Commercial Type and Domain Type, $F(1, 58) = 20.08, p < .001$.

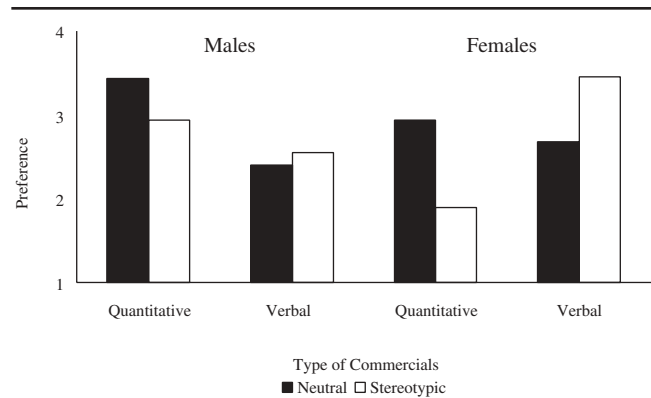


Figure 2 Preferences for majors/careers in quantitative/verbal domains as a function of gender and commercial type.

Simple-effect tests on the women's data revealed a tendency for women exposed to the neutral commercials, similar to men in either condition, to express more interest in the Quantitative Domain ($M = 2.94$) than the Verbal Domain ($M = 2.68$), although this tendency was not significant, $F < 2$. In marked contrast, women exposed to the gender-stereotypic commercials expressed significantly more interest in the Verbal Domain ($M = 3.45$) than the Quantitative Domain ($M = 1.89$), $F(1, 58) = 29.62, p < .001$. Furthermore, simple-effect tests also revealed that women in the neutral-commercial condition expressed significantly more interest in the Quantitative Domain than women in the stereotypic-commercial condition, $F(1, 58) = 6.67, p = .01$. This pattern of interest was reversed for the Verbal Domain, where women in the stereotypic-commercial condition expressed marginally more interest than women in the neutral-commercial condition, $F(1, 58) = 3.67, p = .057$ (see Figure 2).

Study 3 extended our previous research by confirming that stereotype threat can also temporarily affect women's educational and vocational aspirations. Women exposed to gender-stereotypic commercials subsequently indicated less interest in both academic and professional domains in which they risk being negatively stereotyped and more interest in those domains in which they are immune to stereotype threat. Moreover, unlike Study 2, in Study 3, expressing a reduced interest in one field did not necessitate expressing a compensatory interest in another field. Therefore, Study 3 established that stereotype threat leads women not only to avoid traditionally masculine domains but also to actively seek domains in which stereotypes do not allege a sex-based inability.

GENERAL DISCUSSION

Situations in which women risk being personally reduced to a negative stereotype can elicit a disruptive state that interferes with their performance and plea-

sure in those domains. We call this situational predicament stereotype threat. Women are susceptible to stereotype threat whenever they risk being judged by, or treated in terms of, negative stereotypes that provide a plausible explanation for their behavior in a given domain (Steele, 1997). The present research examined the detrimental effects that stereotype threat can have on women's academic performance and achievement-related choices.

In three separate studies, gender-stereotypic television commercials were employed to elicit the female stereotype. Study 1 confirmed that exposure to the stereotypic commercials resulted in activation of the female stereotype among both men and women. But only women, for whom the activated stereotype was self-relevant, underperformed on a subsequent nondiagnostic math test. It also was confirmed that level of stereotype activation among the female participants mediated the performance-inhibiting effects of those television commercials. Study 2 revealed that only women exposed to the gender-stereotypic commercials avoided math items in favor of nonthreatening verbal items on a subsequent nondiagnostic aptitude test. In Study 3, exposure to the stereotypic commercials led women to indicate less interest in educational/vocational options in which they risk being negatively stereotyped (i.e., quantitative domains) and more interest in fields in which stereotypes do not allege a sex-based inability (i.e., verbal domains).

One alternative explanation for our findings is that female participants were simply modeling the behavior of the actresses in the gender-stereotypic commercials. Bandura's social learning theory suggests that media models do indeed have the power to influence the viewer's behavior (e.g., Bandura, Ross, & Ross, 1961). Directly modeling the actresses' behavior in these commercials, however, cannot fully account for the female participants' performance on the aptitude test in Study 2 or their educational/vocational aspirations in Study 3. The actresses do behave in a generally unintelligent manner, but if female viewers were simply modeling unintelligent behavior, they would have underperformed on the entire aptitude test and shied away from all intellectually demanding majors/careers. As reported earlier, however, women exposed to the gender-stereotypic commercials underperformed only on the math portion of the aptitude test and shied away from only quantitative domains.

Fredrickson's work on self-objectification (e.g., Fredrickson, Roberts, Noll, Quinn, & Twenge, 1998) provides another potential counterexplanation for our findings. Objectification theory proposes that our culture objectifies women's bodies, which encourages women to take a third-person perspective of their

physical selves. Fredrickson suggests that this self-objectification not only produces body shame among women but also consumes attentional resources that can manifest as diminished mental performance. Circumstances that elicit a state of self-objectification can therefore lead to underperformance on any cognitively demanding activity. It is possible that viewing our gender-stereotypic commercials could elicit a state of self-objectification, but that still would not explain our results in Study 2. As Fredrickson has argued, self-objectification would disrupt not only women's math performance but also their verbal performance because both are cognitively demanding tasks.

One could argue that the commercials employed in the present research activated different scripts for the male and female participants, and only women were primed to behave in a stereotype-consistent manner. Challenging this alternative explanation is the fact that women experiencing stereotype threat in Study 3 sought domains that were not stereotypically feminine (e.g., political science, novelist, journalist, editor, etc.). This script counterexplanation is also challenged by our research showing the stifling effect that stereotype threat can have on women's leadership aspirations (Davies & Spencer, 2002). Several studies have revealed that women exposed to our gender-stereotypic commercials normally avoid leadership roles in favor of subordinate roles on an impending leadership task. Once again, we found that exposure to the stereotypic commercials primed the female stereotype among participants; however, activation level only predicted role preference among women vulnerable to stereotype threat; that is, regardless of their level of stereotype activation, when women did not risk being personally reduced to a negative stereotype their interest in leadership was restored. Varying the stereotype relevance of the leadership task moderated the effect of the commercials on women's role preference (Davies & Spencer, 2002). If our gender-stereotypic commercials were simply priming women to behave in a stereotype-consistent manner, removing stereotype threat from the impending task would not have restored women's interest in the traditionally masculine domain of leadership.

The ideomotor perspective provides another potential counterexplanation for our findings. As discussed earlier, ideomotor theory suggests that primed stereotypes lead automatically and unconsciously to stereotype-consistent behavior, regardless of the stereotype's relevance to the target. Study 1 confirmed that exposure to the gender-stereotypic commercials primed the female stereotype among both men and women; however, only participants for whom the primed stereotype was self-relevant (i.e., women) underperformed on the subse-

quent math test. Stereotype relevance moderated whether activation level mediated our stereotype-threat effects. In fact, participants for whom the primed stereotype was not self-relevant (i.e., men) actually revealed a consistent tendency for improved performance on subsequent math tests, a phenomenon termed “stereotype lift” (see Walton & Cohen, 2002). In a meta-analytic review, data from numerous studies revealed this reliable performance boost among nonstigmatized participants aware of negative stereotypes targeted at stigmatized others (Walton & Cohen, 2002). Consequently, we assume that in most real-life situations an individual’s motivation to do well should override whatever suggestive influence an activated, but personally irrelevant, stereotype may provide (see Steele et al., in press). For example, the classic Bargh et al. (1996) ideomotor effect, which was discussed earlier in this article, should be overridden by motivation when the undergraduates believe it is a race to the elevator. Some of the inconsistencies between stereotype-threat and ideomotor findings may be explained by examining this relationship between personal motivation and the suggestive influence of personally irrelevant stereotypes. We fully agree with ideomotor researchers when they suggest that primed stereotypes will sometimes influence behavior through cold processes (i.e., ideomotor perspective), sometimes through hot processes (i.e., stereotype-threat perspective), and sometimes through a combination of hot and cold processes (see Wheeler & Petty, 2001).

Individuals base their achievement-related choices on perceptions of how interesting and appropriate a certain domain appears, their perceived success in that domain, and their likelihood of being stereotyped in that domain (Stangor & Sechrist, 1998). We believe that stereotype threat can potentially taint all of these individual perceptions and consequently govern women’s achievement-related choices. The discovery that stereotype threat can lead to the avoidance of targeted domains may represent the seeds of psychological disengagement. Members of stigmatized groups may try to cope with stereotype threat by temporarily disengaging their self-esteem from threatening domains. This defensive detachment ensures that feelings of self-worth are independent of any performance feedback received in those domains (Major et al., 1998). Similar to stereotype threat, psychological disengagement can be a temporary state that is situation specific, but it normally occurs in situations where poor performance is either experienced or anticipated (Crocker et al., 1998; Major et al., 1998). For instance, Major and her colleagues demonstrated that after the possibility of racial bias was primed, African American students’ self-esteem was not affected by the positive or negative feedback they randomly received on an intelligence test. This finding suggests

that the African American students had at least temporarily disengaged their self-esteem from the stereotype-relevant domain (Major et al., 1998).

As alluded to above, experiencing stereotype threat repeatedly in a given domain may eventually lead from acute to chronic adaptation strategies. That is, temporary psychological disengagement may eventually result in stable disidentification (Crocker et al., 1998; Major et al., 1998; Steele, 1992). Disidentification is an adaptation strategy of eliminating a domain as a long-term basis of self-evaluation. For example, if women repeatedly experience stereotype threat in traditionally masculine domains, to protect their feelings of self-worth, they may permanently disassociate themselves from those domains. Disidentification can therefore lead to systematic group differences in aspirations and accomplishments in stereotype-relevant fields. Inevitably, the stigmatized group’s lack of achievement in those domains is then erroneously interpreted by others as reflecting some group-based inability rather than simply a defensive effort to cope with a threat imposed by their own society (Crocker et al., 1998; Crocker & Major, 1989).

Throughout the present research, women’s response to the gender-stereotypic commercials signified an interaction between their social identity and a situational predicament created by way of cultural stereotypes. Their response was not indicative of any internalized doubts about their ability, or their group’s ability, in a given domain. In fact, all of our female participants indicated on a mass-testing questionnaire that they were good at math and being good at math was important to them. Clearly, these women do not believe the negative accusations that allege a sex-based math inability, yet they still must contend with the threat of being personally viewed through stereotypic lenses. As long as negative stereotypes permeate our mass-media culture, members of stigmatized groups will be susceptible to the insidious effects of stereotype threat. The elimination of negative stereotypes from our society may be an unattainable goal but our research has shown that it is possible to eliminate the situational cues that give rise to stereotype threat. Therefore, a feasible remedy to this predicament is to create environments in which stigmatized individuals can work and study without the threat of being personally reduced to a negative stereotype. Within these identity-safe environments the insidious effects of stereotype threat will be disabled.

NOTES

1. A three-way ANOVA (Gender \times Commercial Type \times Question Type) on the performance data controlling for number of attempts (corrected score/number of attempts) revealed a main effect for question type, $F(1, 63) = 83.39, p < .001$, which was qualified by a two-way interaction between Gender and Question Type, $F(1, 63) = 9.65, p < .01$, which in turn was qualified by a marginal three-way interaction

between Gender, Commercial Type, and Question Type, $F(1, 63) = 3.70$, $p = .06$. The women's data were as follows: math performance in the neutral-commercial condition $M = .53$, math performance in the stereotypic-commercial condition $M = .39$, verbal performance in the neutral-commercial condition $M = .24$, and verbal performance in the stereotypic-commercial condition $M = .28$.

2. None of the interactions involving the Health factor were significant, $F_s < 2$. Therefore, only the cell means will be reported here. The means for men in the neutral-commercial condition, men in the stereotypic-commercial condition, women in the neutral-commercial condition, and women in the stereotypic-commercial condition, respectively, were as follows: $M = 2.81$, $M = 3.11$, $M = 3.98$, $M = 3.27$.

REFERENCES

- Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C., & 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.
- Aronson, J., Quinn, D., & Spencer, S. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83-103). San Diego, CA: Academic Press.
- Ashmore, R. D., Del Boca, F. K., & Wohlers, A. J. (1986). Gender stereotypes. In R. D. Ashmore & F. K. Del Boca (Eds.), *The social psychology of female-male relations* (pp. 69-119). New York: Academic Press.
- Bandura, A., Ross, D., & Ross, S. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575-582.
- 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.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.
- Blascovich, J., Spencer, S. J., Quinn, D. M., & Steele, C. M. (2001). Stereotype threat and the cardiovascular reactivity of African Americans. *Psychological Science*, *12*, 225-229.
- Bloom, A. (1968). *The republic of Plato*. New York: Basic Books.
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, *96*, 608-630.
- Crocker, J., Major, B., & Steele, C. M. (1998). Social stigma. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 504-553). Boston: McGraw-Hill.
- Culotta, E., & Gibbons, A. (1992, November 13). Minorities in science. *Science*, *258*, 1176-1232.
- Davies, P. G. (2000). *Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally*. Unpublished doctoral dissertation, University of Waterloo, Ontario, Canada.
- Davies, P. G., & Spencer, S. J. (2002). *Advertisements that elicit stereotype threat can persuade women to avoid leadership positions*. Presented at the Society for Personality and Social Psychology annual meeting, Savannah, GA.
- Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D., Koomen, W., van Knippenberg, A., & Scheepers, D. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology*, *75*, 862-871.
- Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly*, *18*, 585-609.
- Educational Testing Service. (1987). *Practicing to take the GRE General Exam*. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (1992). *Practicing to take the GRE General Test—No. 9*. Princeton, NJ: Educational Testing Service.
- Fredrickson, B. L., Roberts, T., Noll, S. M., Quinn, D. M., & Twenge, J. M. (1998). That swimsuit becomes you: Sex differences in self-objectification, restrained eating, and math performance. *Journal of Personality and Social Psychology*, *75*, 269-284.
- Gerbner, G., Gross, L., Morgan, M., & Signorielli, N. (1993). Growing up with television: The cultivation perspective. In J. Bryant & D. Zillman (Eds.), *Media effects: Advances in theory and research*. Hillsdale, NJ: Lawrence Erlbaum.
- Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, *55*, 708-713.
- Hewitt, N. M., & Seymour, E. (1991). *Factors contributing to high attrition rates among science and engineering undergraduate majors*. A report submitted to the Alfred P. Sloan Foundation.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*, 602-619.
- Kimball, M. M. (1986). Television and sex-role attitudes. In T. M. Williams (Ed.), *The impact of television: A natural experiment in three communities* (pp. 265-301). Orlando, FL: Academic Press.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology*, *82*, 283-299.
- Lefevre, J., Kulak, A., & Heymans, S. (1992). Factors influencing the selection of university majors varying in mathematical content. *Canadian Journal of Behavioral Sciences*, *24*, 276-289.
- Leyens, J. P., Désert, 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., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24*(1), 34-50.
- Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology*, *26*, 291-310.
- Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview.
- 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.
- Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology*, *75*, 1191-1197.
- Stangor, C., & Sechrist, G. B. (1998). Conceptualizing the determinants of academic choice and task performance across social groups. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 105-124). San Diego, CA: Academic Press.
- Steele, C. M. (1992, April). Race and the schooling of Black Americans. *The Atlantic Monthly*, pp. 68-78.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Steele, C. M., Spencer, S. J., & Aronson, J. (in press). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33). New York: 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.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*, *47A*, 631-650.
- Walton, G. M., & Cohen, G. L. (2002). *Stereotype lift*. Manuscript submitted for publication.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychology Bulletin*, *127*, 797-826.

Received November 6, 2001

Revision accepted February 28, 2002